Comments welcome

Novel Approaches to Coherency Conditions in LDV Models

by
Vassilis Hajivassiliou, Department of Economics, LSE
March 2008

Abstract

The paper discusses the major identification issue of *coherency conditions* in LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. Conditions for coherency as discussed in the existing literature are reviewed and shown to be rather esoteric. Two novel methods for establishing coherency conditions are presented, which have intuitive interpretations and are easy to implement and generalize. The constructive consequence of the new approaches is that they indicate how to achieve coherency in models traditionally classified as incoherent through the use of prior sign restrictions on model parameters. This allows us to develop estimation strategies based on Conditional MLE for simultaneous LDV models without imposing recursivity. Econometric applications are used to illustrate the methods in practice and extensions are given to simultaneous ordered probit models with multiple regions.

A set of extensive Monte-Carlo experiments are used to evaluate the properties of the proposed Conditional MLE and the consequences of employing estimators that make overly restrictive coherency assumptions about the DGP. These experiments confirm very substantive improvements in terms of estimation Mean-Squared-Error by employing the CMLE developed in this paper. They also show that estimators based on the Linear Probability approximation perform poorly in this context.

Our CMLE approach allows for the first time to obtain estimates of the reverse as well as direct interaction terms in LDV models with simultaneity.

Keywords: Limited Dependent Variable Models, Coherency Conditions, Joint Bivariate Probit Model, Linear Probability Models

JEL Classifications: C51, C52, C15

*I have benefited from constructive comments by participants at seminars at the London School of Economics, Cemmap/UCL, CERGE, Prague, and the University of Toulouse, as well as at the 2006 Conference on Panel Data at the University of Cambridge and the 2006 Corporate Finance Conference in London. I thank in particular Richard Blundell, Guy Laroque, Peter Robinson, Alain Trognon, and Frank Windjmeier for constructive comments, and Ryan Giordano for expert research assistance. The usual disclaimer applies for any remaining errors.

Address correspondence to:
E-mail: vassilis@lse.ac.uk
Address: Department of Economics
London School of Economics
London WC2A 2AE, England

Novel Approaches to Coherency Conditions in LDV Models

1 Introduction

The paper discusses the major identification issue of *coherency conditions* in LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. The econometric framework of LDV models with simultaneity is presented in Section 2. In the same section we explain the identification issue of *coherency* in such LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables.

Conditions for coherency as discussed in the existing literature are reviewed in Section 6 and shown be rather esoteric. Two novel methods for establishing coherency conditions are presented, one based on a graphical characterization, the second through hypothetical Monte-Carlo DGP. The novel approaches have intuitive interpretations and are easy to implement and generalize. The constructive consequence of the new approaches is that they indicate how to achieve coherency in models traditionally classified as incoherent through the use of prior sign restrictions on model parameters. This allows us to develop estimation strategies in section 7 based on Conditional MLE for simultaneous LDV models without imposing recursivity. Thus one can obtain for the first time estimates of direct as well as reverse interaction effects in simultaneous LDV models, unlike in the existing literature where recursivity had to be assumed. Econometric applications are used to illustrate the methods in practice and extensions are given to simultaneous ordered probit models with multiple regions. Our CMLE approach allows for the first time to obtain estimates of the reverse as well as direct interaction terms in LDV models with simultaneity.

The proposed Conditional MLE methodology is evaluated through an extensive set of Monte-Carlo experiments described in Section 9. The experiments allow us also to study the consequences of employing estimators that make overly restrictive coherency assumptions about the DGP. The findings confirm very substantive improvements in terms of estimation Mean-Squared-Error by employing the CMLE developed in this paper. They also show that estimators based on the Linear Probability approximation perform poorly in this context. Section 11 concludes.

2 The Econometric Problem of “Coherency” in LDV Models

In this section we present and study the fundamental identification issue of *coherency* of LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. Conditions for coherency as discussed in the existing literature are reviewed and shown to be rather esoteric. Two novel methods for establishing coherency conditions are presented, which have intuitive interpretations. Alternative approaches for achieving coherency in models traditionally classified as incoherent through the use of prior restrictions on model parameters.

3 The General Simultaneous LDV Model with Two Interactive Responses

Consider the general two-equation LDV model where limited dependent variables y_1 and y_2 are jointly determined through filter functions $\tau_1(\cdot)$ and $\tau_2(\cdot)$ operating on latent variables y_1^* and y_2^* respectively:

$$y_{1it} = \tau_1 \left(y_{1it}^* \equiv [h_1(x_{1it}'\beta_1, y_{2it}\gamma) + \epsilon_{1it}] \right) \tag{1}$$

$$y_{2it} = \tau_2 \left(y_{2it}^* \equiv [h_2(x_{2it}'\beta_2, y_{1it}\delta) + \epsilon_{2it}] \right) \tag{2}$$

The (possibly non-linear) functions $h_1(\cdot)$ and $h_2(\cdot)$ are known up to parameter vectors β_1 and β_2 and the two interaction coefficients γ and δ. The interaction terms $y_{2it}\gamma$ and $y_{1it}\delta$ appear in the respective latent variables y_{1it}^* and y_{2it}^*. The parameter vector to be estimated is $\theta \equiv (\beta_1', \beta_1', \gamma, \delta, \sigma_1^2, \sigma_2^2, \rho)'$ where $\rho \equiv correlation(\epsilon_{1it}, \epsilon_{2it})$. In the most general case, the sample is a panel data set indexed by $i = 1, \cdots, N$ and $t = 1, \cdots, T$.

The typical *coherency* condition in such models, necessary for the joint distribution $(y_{1it}, y_{2it}|x_1, x_2, \theta)$ to be well-specified is: $\gamma \cdot \delta = 0$. Gourieroux, Laffont, and Monfort (1980)[5] explain condition in terms of there being a valid function from $(\epsilon_{1it}, \epsilon_{2it})$ to the observable endogenous variables (y_{1it}, y_{2it}). Lewbel (2007)[13] establishes NASC for coherency by approaching problem as requiring a valid reduced form system for (y_{1it}, y_{2it}). For example, if $\delta = 0$ then the RF for y_{2it} is:

$$y_{2it} = \tau_2 \left(h_2(x_{2it}'\beta_2) + \epsilon_{2it} \right)$$

and hence the RF for y_{1it} is given by:

$$y_{1it} = \tau_1 \left(h_1(x_{1it}'\beta_2, \gamma \cdot \tau_2 \left(h_2(x_{2it}'\beta_2) + \epsilon_{2it} \right) + \epsilon_{1it} \right)$$

3.1 General Explanation and Illustrative Applications

The leading case we focus on here is the binary threshold crossing response model defined by:

$$\tau_j(z) \equiv \mathbf{1}(z > 0)$$

In terms of the two latent variables y_1^* and y_2^* and the observed binary indicators y_1 and y_2, and suppressing the observation indices:

$$y_1 = \begin{cases} 1 & if \quad y_1^* \equiv x_1\beta_1 + \gamma y_2 + \epsilon_1 > 0 \\ 0 & if \quad y_1^* \equiv x_1\beta_1 + \gamma y_2 + \epsilon_1 \leq 0 \end{cases} \tag{3}$$

$$y_2 = \begin{cases} 1 & if \quad y_2^* \equiv x_2\beta_2 + \delta y_1 + \epsilon_2 > 0 \\ 0 & if \quad y_2^* \equiv x_2\beta_2 + \delta y_1 + \epsilon_2 \leq 0 \end{cases} \tag{4}$$

For a typical i observation, the probability $Prob(y_{1it}, y_{2it}|X, \theta)$ is characterized by the constraints on the unobservables:

$$(a_1, a_2)' < (\epsilon_1, \epsilon_2)' < (b_1, b_2)'$$

through the configuration:

y_{1it}	y_{2it}	a_1	b_1	a_2	b_2
1	1	$-x_{1it}\beta_1 - \gamma$	∞	$-x_{2it}\beta_2 - \delta$	∞
1	0	$-x_{1it}\beta_1$	∞	$-\infty$	$-x_{2it}\beta_2 - \delta$
0	1	$-\infty$	$-x_{1it}\beta_1 - \gamma$	$-x_{2it}\beta_2$	∞
0	0	$-\infty$	$-x_{1it}\beta_1$	$-\infty$	$-x_{2it}\beta_2$

In this case, $(y_1, y_2) \in \{(1,1), (1,0), (0,1), (0,0)\}$ such that:

or:

(y_{1it}, y_{2it})	y_{1it}^*		y_{2it}^*
$(1,1)$	$x_{1it}'\beta_1 + \gamma + \epsilon_{1it} > 0$	,	$x_{2it}'\beta_2 + \delta + \epsilon_{2it} > 0$
$(1,0)$	$x_{1it}'\beta_1 + \epsilon_{1it} > 0$	,	$x_{2it}'\beta_2 + \delta + \epsilon_{2it} < 0$
$(0,1)$	$x_{1it}'\beta_1 + \gamma + \epsilon_{1it} < 0$	,	$x_{2it}'\beta_2 + \epsilon_{2it} > 0$
$(0,0)$	$x_{1it}'\beta_1 + \epsilon_{1it} < 0$	,	$x_{2it}'\beta_2 + \epsilon_{2it} < 0$

In general, in the absence of coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_{1it} + x_{1it}'\beta_1, \epsilon_{2it} + x_{2it}'\beta_2)$.

4 Three Illustrative Models

These models, recently proposed in the literature, did not offer a detailed analysis of their coherency. In this section we will use these models to illustrate how our methods can be implemented in practice.

4.1 Illustrative Model 1: Simultaneous Determination of a Binary Indicator and a Trinomial Ordered Indicator

Let us use a slightly more complicated simultaneous LDV model to illustrate the issue of coherency, namely the *binary & trinomial ordered probit model* of Hajivassiliou and Ioannides (2007)[7] that studies interactions between liquidity and employment constraints on individual households indexed by i at a given point in time indexed by t. Define two latent dependent variables y^*_{1it} and y^*_{2it} and drop the it subscripts:

$$S = \begin{cases} 1 \quad if \quad y^*_1 > 0 \ \text{(liquidity constraint binding)}, \\ 0 \quad if \quad y^*_1 \leq 0 \ \text{(liquidity constraint not binding)}. \end{cases} \tag{5}$$

$$E = \begin{cases} -1 \quad if \quad y^*_2 \leq \lambda^- \ \text{(overemployed)} \\ 0 \quad if \quad \lambda^- \leq y^*_2 < \lambda^+ \ \text{(voluntarily employed)} \\ +1 \quad if \quad \lambda^+ \leq y^*_2 \ \text{(under-/unemployed)}. \end{cases} \tag{6}$$

$$y^*_1 = \mathbf{1}(y^*_2 < \lambda^-)\gamma_{11} + \mathbf{1}(\lambda^- < y^*_2 < \lambda^+)\gamma_{12} + x'_1\beta_1 + \epsilon_1$$

$$y^*_2 = \mathbf{1}(y^*_1 > 0)\delta + x_2\beta_2 + \epsilon_2$$

Since (S, E) lie in $\{0,1\} \times \{-1,0,1\}$, the 6 possible configurations may be enumerated as follows:

S	E	y^*_1	y^*_2
0	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 < 0,$	$x_2\beta_2 + \epsilon_2 < \lambda^-$
0	0	$x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^- < x_2\beta_2 + \epsilon_2 < \lambda^+$
0	1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 < 0,$	$\lambda^+ < x_2\beta_2 + \epsilon_2$
1	-1	$\gamma_{11} + x_1\beta_1 + \epsilon_1 > 0,$	$\delta + x_2\beta_2 + \epsilon_2 < \lambda^-$
1	0	$x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^- < \delta + x_2\beta_2 + \epsilon_2 < \lambda^+$
1	1	$\gamma_{12} + x_1\beta_1 + \epsilon_1 > 0,$	$\lambda^+ < \delta + x_2\beta_2 + \epsilon_2$

In terms of the unobservables, the probability of a (y_1, y_2) observed pair is equivalent to the probability:

$$(a_1, a_2)' < (\epsilon_1, \epsilon_2)' < (b_1, b_2)'$$

where $(\epsilon_1, \epsilon_2)' \sim N(0, \Sigma_\epsilon)$, and a and b are given by:

S	E	a_1	a_2	b_1	b_2
0	-1	$-\infty$	$-\infty$	$-(\gamma_{11} + x_1\beta_1)$	$\lambda^- - x_2\beta_2$
0	0	$-\infty$	$\lambda^- - x_2\beta_2$	$-x_1\beta_1$	$\lambda^+ - x_2\beta_2$
0	1	$-\infty$	$\lambda^+ - x_2\beta_2$	$-(\gamma_{12} + x_1\beta_1)$	$+\infty$
1	-1	$-(\gamma_{11} + x_1\beta_1)$	$-\infty$	$+\infty$	$\lambda^- - \delta - x_2\beta_2$
1	0	$-x_1\beta_1$	$\lambda^- - \delta - x_2\beta_2$	$+\infty$	$\lambda^+ - \delta - x_2\beta_2$
1	1	$-(\gamma_{12} + x_1\beta_1)$	$\lambda^+ - \delta - x_2\beta_2$	$+\infty$	$+\infty$

The variance-covariance matrix captures the contemporaneous correlation between ϵ_1 and ϵ_2. Given the binary nature of S, σ_{11} is normalized to 1. Subsection 8.1 below discusses how to specify this contemporaneous correlation *as well as* flexible forms of serial correlation in panel data settings.

4.2 Illustrative Model 2: Simultaneous Determination of Two Binary Indicators with Observable Dynamics

Next we consider the Currency and Banking Crises model of External Financing of Falcetti and Tudela (2007)[4], to serve as an illustration of how to implement our coherency approach.

Define two latent dependent variables C_{it}^* and B_{it}^* and two binary limited dependent variables C_{it} and B_{it} as follows:

$$C_{it} = \begin{cases} 1 & if C_{it}^* \equiv x_{it}^C \beta^C + \mathbf{1}\left(\sum_{s=1}^4 B_{i,t-s} > 0\right)\zeta^C + B_{it} \cdot \gamma + \epsilon_{it}^C > 0, \\ 0 & otherwise \end{cases} \tag{7}$$

$$B_{it} = \begin{cases} 1 & if B_{it}^* \equiv x_{it}^B \beta^B + +\mathbf{1}\left(\sum_{s=1}^4 C_{i,t-s} > 0\right)\zeta^B + C_{it} \cdot \delta + \epsilon_{it}^B > 0, \\ 0 & otherwise, \end{cases} \tag{8}$$

where $\mathbf{1}(\cdot)$ is the usual indicator function. The dummy dependent variable C_{it} illustrates the occurrence of a currency crisis for country i in period t, while the dummy B_{it} indicates a domestic banking crisis for the country in that time period.

Consider the probability expression: $Prob(C_{i5}, B_{i5}, \cdots, C_{iT_i}, B_{iT_i} | x_i, z_i, C_{i1}, \cdots, C_{i4}, B_{i1}, \cdots, B_{i4}, \theta)$. For a typical observation it:

(C, B)	$C_{it}^* > 0$	B_{it}^*
(1,1)	$\epsilon_{it}^C + x_{it}^C \beta^C + \mathbf{1}_B(\cdot)\zeta^C + \gamma > 0$	$\epsilon_{it}^B + x_{it}^C \beta^C + \mathbf{1}_C(\cdot)\zeta^B + \delta > 0$
(1,0)	$\epsilon_{it}^C + x_{it}^C \beta^C + \mathbf{1}_B(\cdot)\zeta^C > 0$	$\epsilon_{it}^B + x_{it}^C \beta^C + \mathbf{1}_C(\cdot)\zeta^B + \delta > 0$
(0,1)	$\epsilon_{it}^C + x_{it}^C \beta^C + \mathbf{1}_B(\cdot)\zeta^C + \gamma > 0$	$\epsilon_{it}^B + x_{it}^C \beta^C + \mathbf{1}_C(\cdot)\zeta^B > 0$
(0,0)	$\epsilon_{it}^C + x_{it}^C \beta^C + \mathbf{1}_B(\cdot)\zeta^C > 0$	$\epsilon_{it}^B + x_{it}^C \beta^C + \mathbf{1}_C(\cdot)\zeta^B > 0$

and in terms of constraints on the unobservables, when writing $\mu^B = x^B \beta^B$, $\mu^C = x^C \beta^C$:

(C,B)	a^C	a^B	b^C	b^B
(1,1)	$-\mu^C-1_B(\cdot)\zeta^C-\gamma$	$-\mu^B-1_C(\cdot)\zeta^B-\delta$	∞	
(1,0)	$-\mu^C-1_B(\cdot)\zeta^C$	$-\infty$	∞	$-\mu^B-1_C(\cdot)\zeta^B-\delta$
(0,1)	$-\infty$	$-\mu^B-1_C(\cdot)\zeta^B-\delta$	$-\mu^C-1_B(\cdot)\zeta^C-\gamma$	∞
(0,0)	$-\infty$	$-\infty$	$-\mu^C-1_B(\cdot)\zeta^C$	$-\mu^B-1_C(\cdot)\zeta^B$

4.3 Illustrative Model 3: Simultaneous Determination of Two Binary Indicators

Hajivassiliou and Savignac (2007)[9] use a joint binary probit model to study the impact of financing constraints on a firm's decision and ability to innovate.

Define two latent dependent variables I_{it}^* and F_{it}^* and two binary limited dependent variables I_{it} and F_{it} as follows:

$$I = \begin{cases} 1 & if \quad I^* \equiv x^I\beta^I + \gamma F + \epsilon^I > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$F = \begin{cases} 1 & if \quad F^* \equiv x^F\beta^F + \delta I + \epsilon^F > 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

For a typical it observation, the probability $Prob(I_{it}, F_{it}|X, \theta)$ is characterized by the constraints on the unobservables:

$$(a^I, a^F)' < (\epsilon^I, \epsilon^F)' < (b^I, b^F)'$$

through the configuration:

I_{it}	F_{it}	a^C	b^C	a^B	b^B
1	1	$-x_{it}^I\beta^I - \gamma$	∞	$-x_{it}^F\beta^F - \delta$	∞
1	0	$-x_{it}^I\beta^I$	∞	$-\infty$	$-x_{it}^F\beta^F - \delta$
0	1	$-\infty$	$-x_{it}^I\beta^I - \gamma$	$-x_{it}^F\beta^F$	∞
0	0	$-\infty$	$-x_{it}^I\beta^I$	$-\infty$	$-x_{it}^F\beta^F$

This model corresponds to the leading model we introduced at the beginning of section 3.1.

5 The Importance of Simultaneity and Sample Selection

The findings of Hajivassiliou and Savignac (op.cit.) highlight the possibly huge impact of allowing correctly for simultaneity in joint LDV models, which if ignored or incorrectly specified, can lead to dramatic inconsistencies in the estimated coefficients. In that study, the estimated coefficient of the ceteris paribus impact of a binding financing constraint on a firm's willingness and ability to innovate ranges from -1.29 to over $+0.55$, and reported statistically significant in all estimated versions!

6 The Traditional Approach to Coherency Conditions

To maintain the logical consistency of the model (known in the literature as "coherency") y_1^* should not depend on y_2^* if y_2^* depends on y_1^* and vice-versa. Let us use the slightly more complicated simultaneous LDV model of Hajivassiliou and Ioannides (op.cit.) discussed in Section 4.1. We have seen there that the six possible configurations of the unobservables $(\epsilon_1, \epsilon_2)'$ of the model correspond to:

S	E	a_1	a_2	b_1	b_2
0	-1	$-\infty$	$-\infty$	$-(\gamma_{11} + x_1\beta_1)$	$\lambda^- - x_2\beta_2$
0	0	$-\infty$	$\lambda^- - x_2\beta_2$	$-x_1\beta_1$	$\lambda^+ - x_2\beta_2$
0	+1	$-\infty$	$\lambda^+ - x_2\beta_2$	$-(\gamma_{12} + x_1\beta_1)$	$+\infty$
1	-1	$-(\gamma_{11} + x_1\beta_1)$	$-\infty$	$+\infty$	$\lambda^- - \delta - x_2\beta_2$
1	0	$-x_1\beta_1$	$\lambda^- - \delta - x_2\beta_2$	$+\infty$	$\lambda^+ - \delta - x_2\beta_2$
1	+1	$-(\gamma_{12} + x_1\beta_1)$	$\lambda^+ - \delta - x_2\beta_2$	$+\infty$	$+\infty$

Using traditional arguments, we obtain that a sufficient condition for coherency of the model is: $(\gamma_{11} + \gamma_{12})\delta = 0$ and $\gamma_{11}\gamma_{12}\delta = 0$.

- To verify this condition, suppose $(S, E) = (0, 0)$. This rules out $(S, E) = (0, -1)$ because $x_2\beta_2 + \epsilon_2 > \lambda^-$, and rules out $(S, E) = (1, 0)$ because $x_1\beta_1 + \epsilon_1 < 0$.
- But $(1, -1)$ is not ruled out if the coherency conditions do not hold, since γ_{11} could be sufficiently negative and δ sufficiently positive to imply the $(1, -1)$ conditions.
- Similarly, the $(1, 1)$ possibility cannot be ruled out in the absence of the coherency conditions, since γ_{12} and δ can be sufficiently positive.
- Such logical inconsistencies are prevented if either (a) $\delta = 0$ or (b) γ_{11} and γ_{12} are simultaneously 0.

Similar considerations can be employed to establish that the traditional coherency condition for the joint binary probit models of Models 2 and 3 *while assuming no intertemporal endogeneity or dynamics* are: $\gamma \cdot \delta = 0$.

This condition, of course, translates to the models (7)-(8) and (9)-(10) being *recursive*. See Maddala and Lee (1976)[14].

It is very important to note that in case the joint binary probit model were allowed to contain intertemporal endogeneity of the type contained in (7)-(8) and the dynamic versions of (9)-(10) estimated by Hajivassiliou and Savignac (op.cit.) [reported in their tables 3 and 4), the coherency condition is practically impossible to generalize and verify using the traditional analysis given in the previous paragraph.

6.1 Difficulties with the traditional approaches:

The first difficulty is that derivations of formal conditions using the traditional approach lack intuition. Second, they are practically impossible to generalize and verify in moderately more complicated LDV models, especially in cases where the models are allowed to contain intertemporal endogeneity of the type contained in (7)-(8) and the dynamic versions of (9)-(10)

The third major difficulty that in practice, non-triangular or reverse triangular cases are the most interesting from an economic point of view. Finally, the traditional approaches focus on establishing *necessary* conditions for coherency, which our methods allow us to prove that they are *not sufficient*.

To overcome the first two difficulties, alternative ways for establishing coherency are developed here, that are both intuitive and straightforward, as well as much more generalizable. In addition, our methods allow us to resolve the last two difficulties leading to estimation based on Conditional MLE for much more interesting practical applications. It is shown in the next Section how to establish coherency without recursiveness through the use of (a) endogeneity in terms of latent variables and/or (b) sign restrictions on model parameters. The fact that our novel approach for the first time eliminates the need to assume recursivity is quite important for the economic problem studied in Hajivassiliou and Savignac (op.cit.): recursivity corresponds to the key identifying assumption that innovation does not affect financial distress directly ($\delta = 0$). On a priori grounds, this assumption seems particularly dubious since innovation may lead to more profits and thus relax financial constraints (corresponding to $\delta > 0$). An alternative possibility is that innovation may lead to higher investment in intangible assets thus reinforcing binding financial constraints (corresponding to $\delta < 0$). Both possibilities violate the traditional coherency condition.[1]

6.2 Novel Approach 1: Graphical

Let us illustrate the first approach using the Liquidity-Employment constraints application of Hajivassiliou and Ioannides (op.cit.). It should be noted that this graphical approach is related to that of [16] who studied the problem of coherency in bivariate discrete models for games with multiple equilibria. Figure overleaf 1 gives the 6 possible regimes $(S \times E) = \{1, 0\} \times \{-1, 0, 1\}$ in terms of the two latent variables y_1^* and y_2^* and the possible configurations in terms of parameters $\bar{\lambda}$, $\underline{\lambda}$, δ, γ_{11}, and γ_{12}. y_1^* is on the horizontal axis and y_2^* on the vertical.

The figure makes clear the role of the coherency condition (a) $\delta = 0$ or (b) $\gamma_{11} = \gamma_{12} = 0$: in general, regions $R2$ and $R6$ exhibit double-counting (cross-hatched area), as well as a white rectangle remains which makes the six regions not mutu-

[1] Note that throughout we expect $\gamma < 0$, i.e., the higher the probability that a firm faces a binding financial constraint, the less likely it is that it is able to innovate. So the two possibilities translate to: (a) $\gamma < 0$, $\delta > 0$ and (b) $\gamma < 0$, $\delta < 0$.

ally exhaustive. These two logical incoherencies disappear when either $\delta = 0$ and/or $\gamma_{11} = \gamma_{12} = 0$ hold.

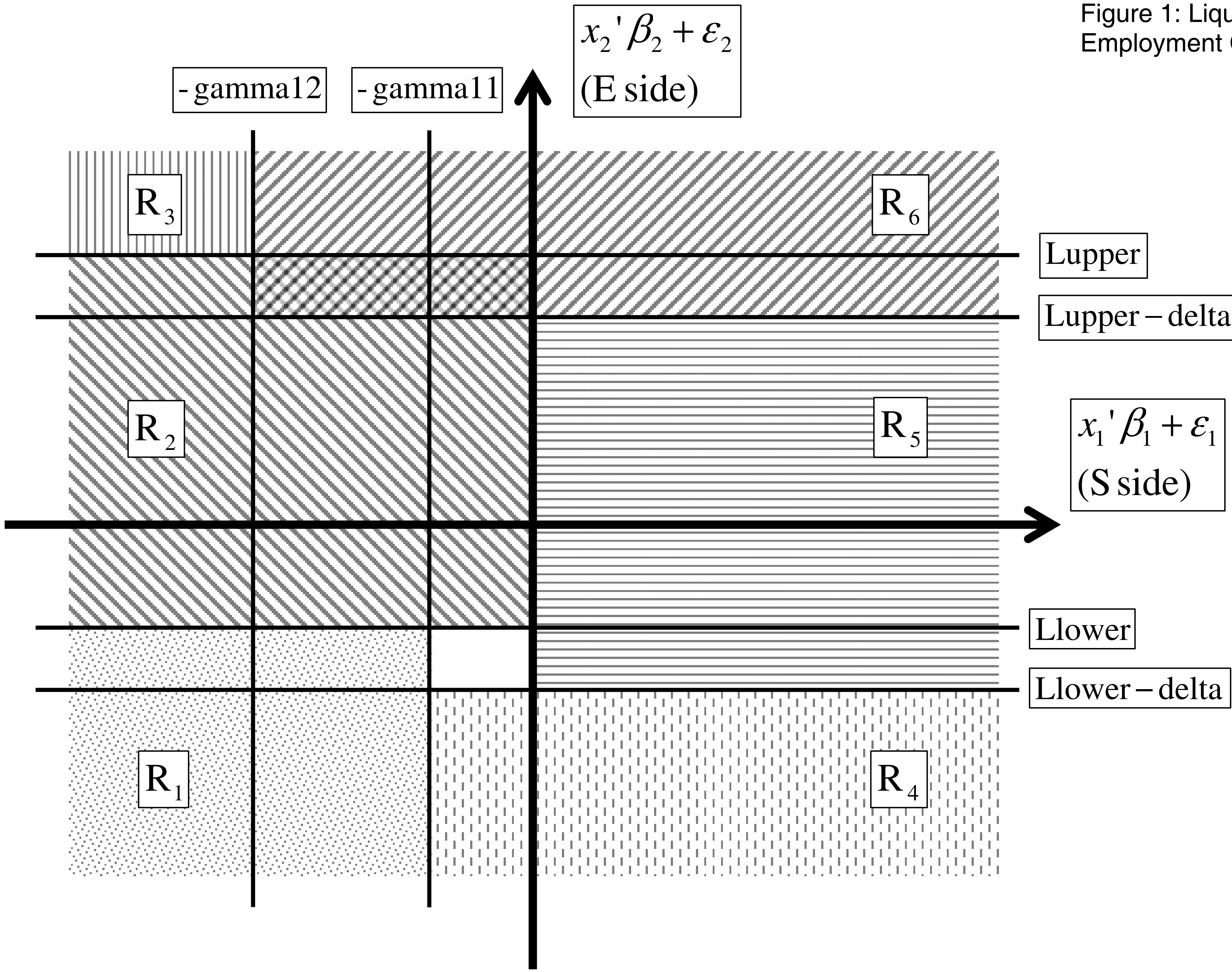
We develop further our graphical approach in Section 7 below, and use it to highlight the fundamental distinction between two types of incoherency, the first corresponding to overlap regions in latent variables space, while the second to empty regions. We explain there that incoherencies of the latter type can be overcome through additional prior restrictions on model parameters through the use of Conditional MLE.

6.3 Novel approach 2: DGP From First Principles

The second approach to incoherency consists of designing a data-generating algorithm (on a computer or hypothetical) to simulate random draws from an LDV model's structure. Again let us use the Liquidity-Employment Constraints application of Hajivassiliou and Ioannides (op.cit.) to illustrate the method. We draw ϵ_1 and ϵ_2 under the joint bivariate normal distribution with zero mean vector and variance-covariance matrix Σ_ϵ, and given $x_1\beta_1$ and $x_2\beta_2$ attempt to generate y_1^* and y_2^*. This is possible provided the coherency condition holds: If (a) $\delta = 0$, then latent y_2^* can be drawn, then ldv y_2, which together with ϵ_1 and $x_1\beta_1$ determines the rhs of y_1^*, thus allowing y_1 to be drawn. Similarly, if (b) $\gamma_{11} = \gamma_{12} = 0$, then y_1^* can be drawn from the first equation based on ϵ_1 and $x_1\beta_1$, which determines y_1, thus giving y_2^* and hence y_2. It is not obvious, however, whether such data generation can be achieved in case the coherency condition does not hold. This approach is related to the Gourieroux et al. (op.cit.) condition that a function exist from ϵ_1, ϵ_2 to y_1, y_2.

As we will show in section 8.1 below, the approach extends naturally to cases with intertemporal endogeneities in panel LDV models, and can be used to prove the coherency of the classic multiperiod panel probit with state dependence (Heckman (1981)[11]), as well as the general versions of illustrative models 2 and 3 with explicit dynamic effects.

Figure 1: Liquidity and Employment Constraints



7 Identification Under Additional Prior Sign Restrictions

The graphical approach we developed in the previous section highlights two distinct cases of incoherency: the first type of incoherency corresponds to regions of the observed endogenous variables of the model being *overlapping*, while the second to regions that are *empty*. It is shown below (a) that overlapping region incoherency can be transformed into empty region incoherency by redefining one of the observed binary LDVs to its complement. And (b) that empty region incoherency can be overcoming through conditional maximum likelihood (CMLE) of truncating the LDVs to lie outside the incoherency regions.

The CMLE approach we propose here can also be motivated through the DGP approach to establishing coherency, that we discussed in the previous subsection. In that case, we need to consider DGPs truncated to lie on a specific region of the latent variables space. A specific method for achieving this is given in technical Appendix 1 below.

It is also useful to note that our approach for establishing coherency through the use of prior sign restrictions developed here is related to the recent approach by Uhlig (2005)[17] for VAR identification under prior sign restrictions on impulse response functions.[2]Dagenais (1997)[3] also makes a distinction between alternative types of incoherency regions.[3]

7.1 Latent Variable Endogeneity

For completeness, let us modify the two-equation LDV model (1)-(2) to make the interaction terms be the *latent* variables instead of the *limited* counterparts:

$$y_{1it} = \tau_1 \left(y_{1it}^* \equiv [h_1(x_{1it}'\beta_1, y_{2it}^*\gamma) + \epsilon_{1it}]\right)$$
$$y_{2it} = \tau_2 \left(y_{2it}^* \equiv [h_2(x_{2it}'\beta_2, y_{1it}^*\delta) + \epsilon_{2it}]\right)$$

Then:

$$y_1^* = x_1\beta_1 + y_2^*\gamma + \epsilon_1$$
$$y_2^* = x_2\beta_2 + y_1^*\delta + \epsilon_2$$

and

$$y_1^* = x_1\beta_1 + \gamma \cdot [x_2\beta_2 + y_1^*\delta + \epsilon_2] + \epsilon_1$$
$$y_2^* = x_2\beta_2 + \delta \cdot [x_1\beta_1 + y_2^*\gamma + \epsilon_1] + \epsilon_2$$

[2]I am indebted to Alain Trognon for pointing out the potential of parameter sign restrictions overcoming incoherency of the "empty region" type, and to Hashem Pesaran for bringing to my attention Uhlig's work on sign identification.

[3]Unfortunately his work remains incomplete and unpublished due to his untimely death.

Hence $y_1^* = RF_1$ and $y_2^* = RF_2$, allowing us to obtain $y_1 = \tau(RF_1)$ and $y_2 = \tau(RF_2)$. We thus see that it is considerably more straightforward to establish coherency identification of LDV models with latent variable interactions as opposed to limited variable interactions.

7.2 Coherency through Sign Restrictions

We illustrate the Conditional MLE approach using the joint binary probit model:[4]

$$I = \begin{cases} 1 & if\_I^* \equiv x^I \beta^I + \gamma F + \epsilon^I > 0 \\ 0 & otherwise \end{cases}$$

$$F = \begin{cases} 1 & if\_F^* \equiv x^F \beta^F + \delta I + \epsilon^F > 0 \\ 0 & otherwise \end{cases}$$

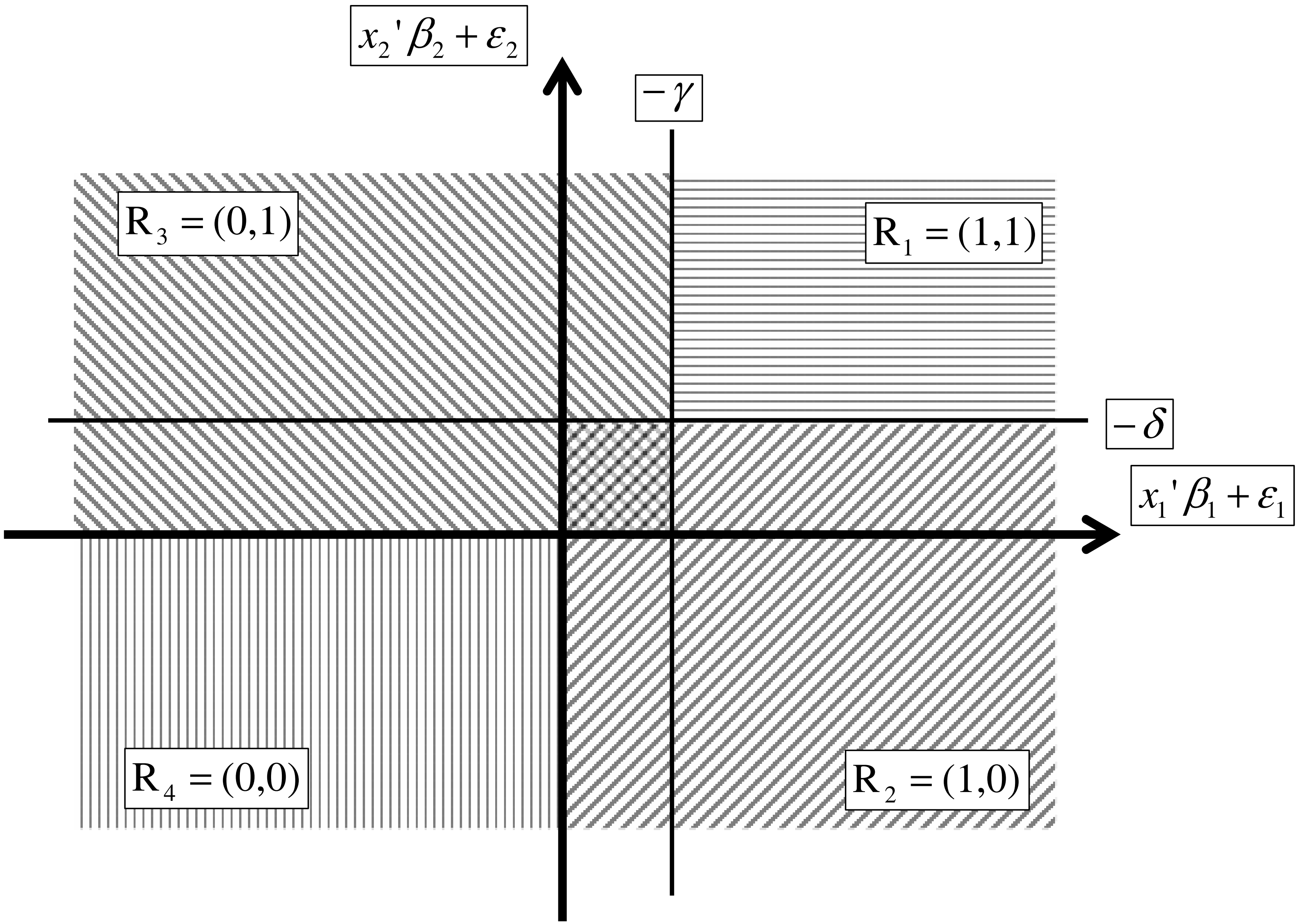
Obviously, there exist **four cases** based on signs of γ, δ:

[4]For the first equation, I^* is used for the latent and I for the observed LDV as a mnemonic to the *Innovation* side of the model of Hajivassiliou and Savignac (2007). Similarly, for the second equation we use F^* and F as a mnemonic to *Financing Constraints*.

7.3 Case 1: $\gamma > 0, \delta > 0$ — overlapping regions, incoherency

The Conditional MLE methodology is not directly applicable to this case:

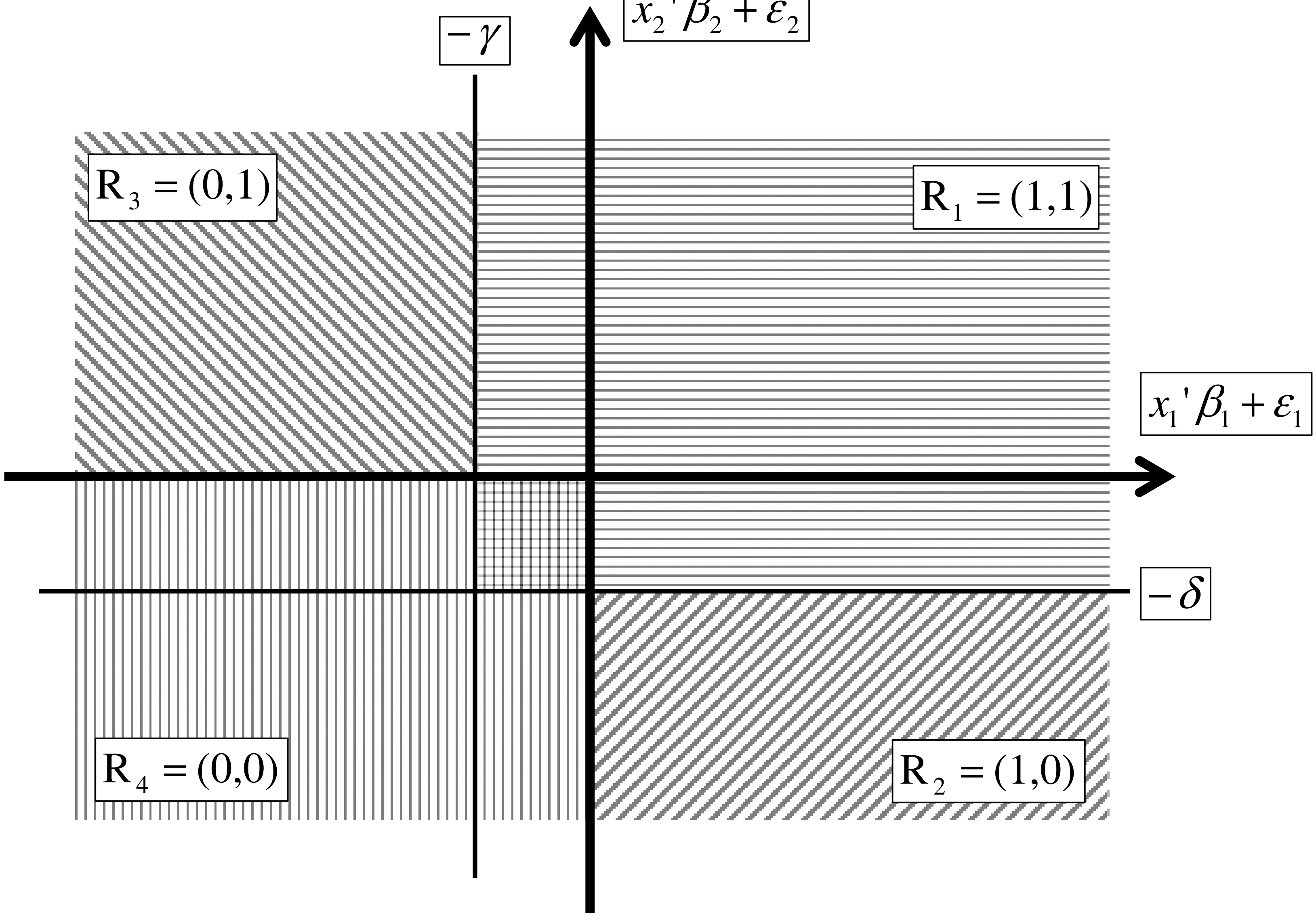
Figure 2: Case 1 (g>0,d>0)



7.4 Case 2: $\gamma < 0, \delta < 0$ — overlapping regions, incoherency

The Conditional MLE methodology is not directly applicable to this case either:

Figure 3: Case 2 (g<0,d<0)

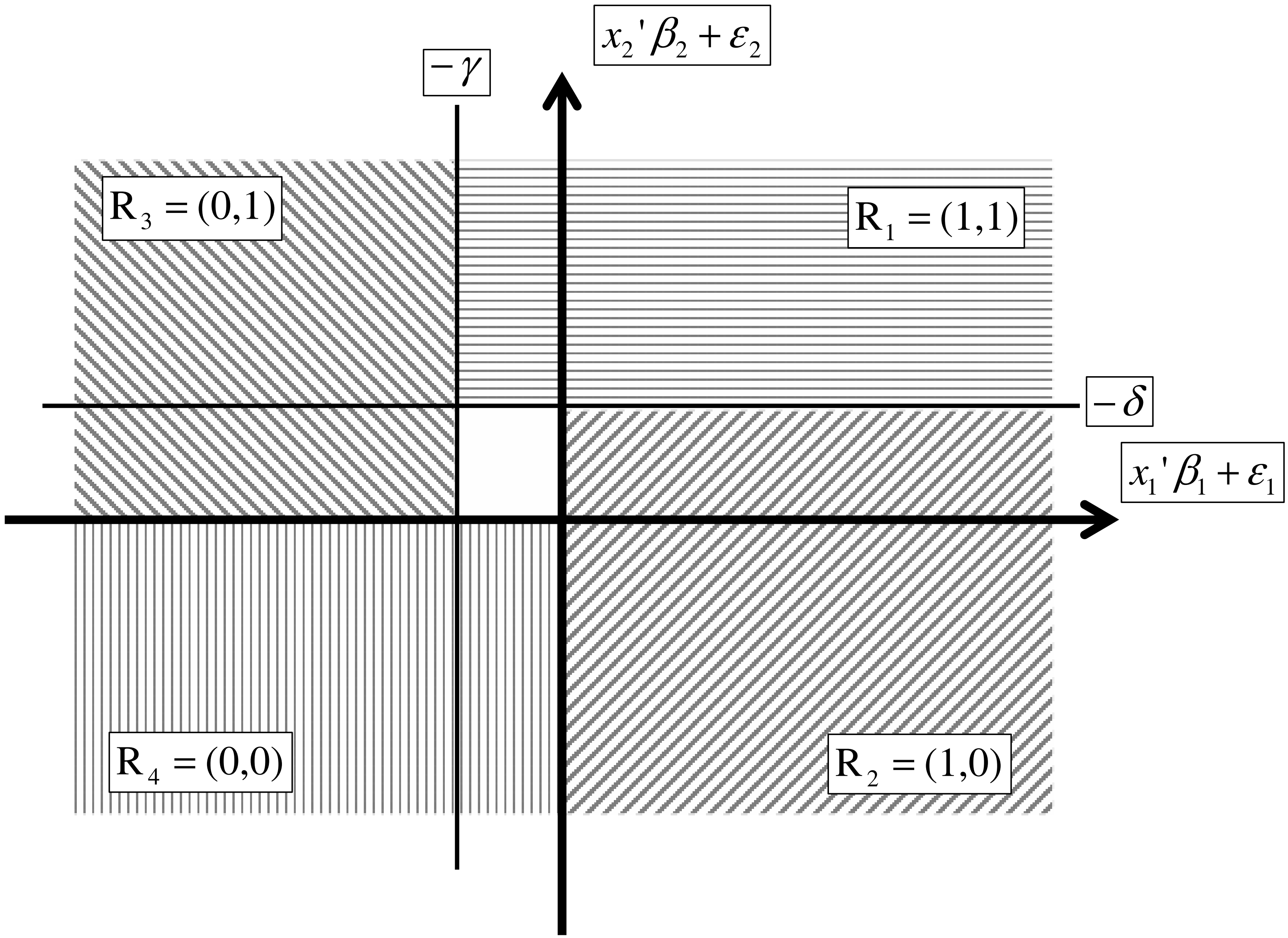


7.5 Case 3: $\gamma > 0, \delta < 0$ — empty regions, coherency through conditioning

For this case, coherency can be achieved using CMLE by conditioning the observed LDVs to lie outside the "empty" region of figure 4, which has conditioning probability:

$$1 - Prob(-\gamma < \epsilon_1 + x_1\beta_1 < 0, 0 < \epsilon_2 + x_2\beta_2 < -\delta)$$

Figure 4: Case 3 (g>0,d<0)

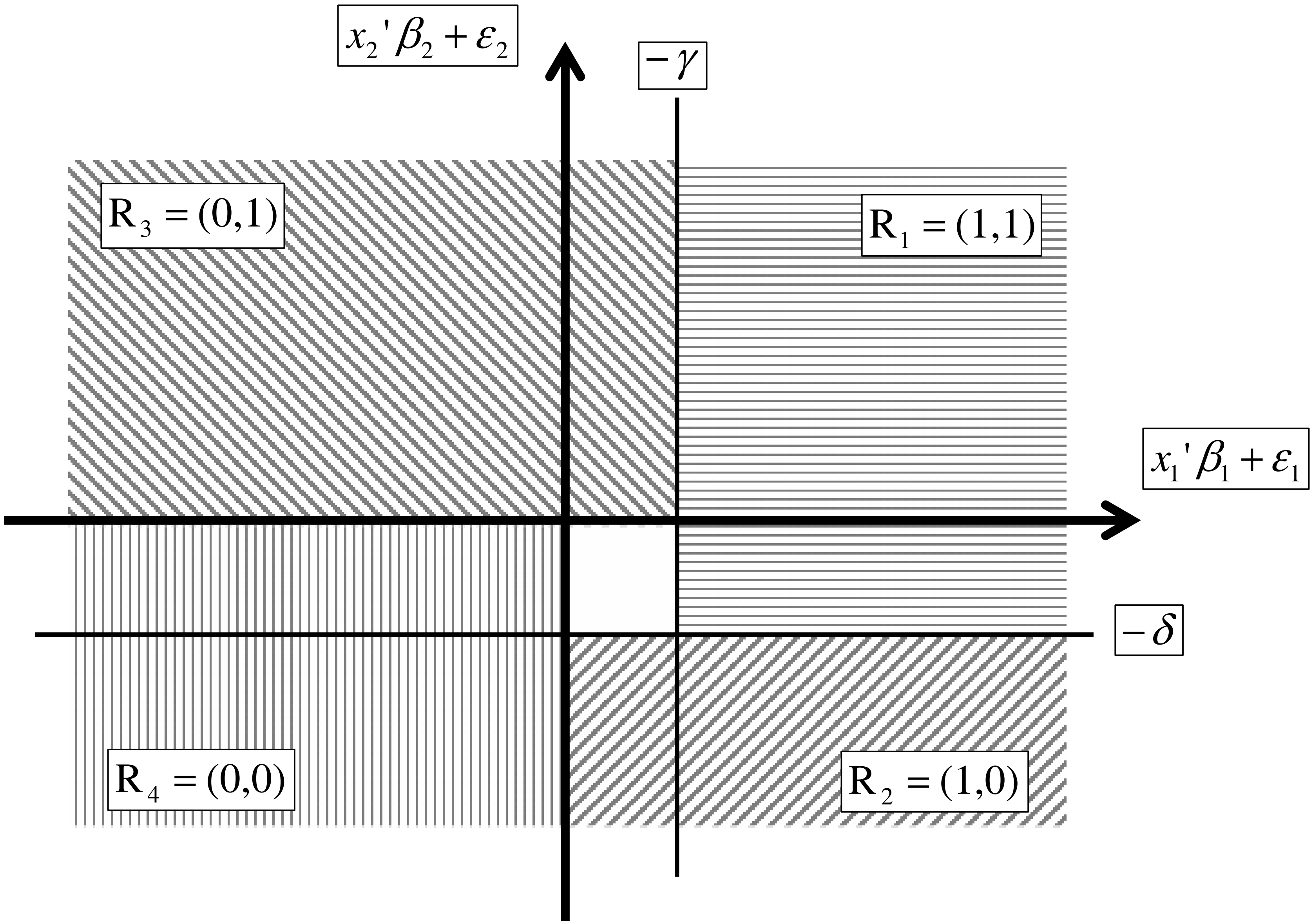


7.6 Case 4: $\gamma < 0, \delta > 0$ — empty regions, coherency through conditioning

For this case, coherency is also achievable through CMLE by conditioning to the LDVs to lie outside the "empty" region of figure 5, which has conditioning probability:

$$1 - Prob(0 < \epsilon_1 + x_1\beta_1 < -\gamma, \delta < \epsilon_2 + x_2\beta_2 < 0)$$

Figure 5: Case 4 (g>0,d<0)



7.7 To Show that Models with Overlapping Regions Remain Incoherent Irrespective of LDV Definitions

We have shown that in general, in the absence of coherency conditions, there will be *overlaps* and/or *gaps* in the domain of $(\epsilon_1 + x_1'\beta_1, \epsilon_2 + x_2'\beta_2)$. At this point, a researcher might be tempted to propose that the incoherency cases with overlapping regions (Cases 1 and 2 above) may be overcome by redefining one of the two limited dependent variables to their complement. According to this reasoning, since the incoherency is caused in these cases because γ and δ are of the same sign, and since changing y_2, say, to its complement $y_2^N \equiv (1 - y_2)$ would result in $\delta^N \equiv -\delta$, then coherency would be achieved since then $\gamma \cdot \delta^N < 0$.

Such reasoning would be incorrect, however. We analyze here this idea and show that such a redefinition would *maintain* the overlapping-region incoherency. This is because the $y_2^N \equiv (1 - y_2)$ redefinition would also switch the sign of γ and hence $\gamma^N \cdot \delta^N > 0$ just as $\gamma \cdot \delta > 0$.

Let us return to the bivariate binomial probit model in terms of the two latent variables I^* and FC^* and the observed binary indicators I and FC, and suppressing the observation index:

$$I = \begin{cases} 1 & if \quad I^* \equiv x_1\beta_1 + \gamma FC + \epsilon_1 > 0 \\ 0 & if \quad I^* \equiv x_1\beta_1 + \gamma FC + \epsilon_1 \leq 0 \end{cases} \tag{11}$$

$$FC = \begin{cases} 1 & if \quad FC^* \equiv x_2\beta_2 + \delta I + \epsilon_2 > 0 \\ 0 & if \quad FC^* \equiv x_2\beta_2 + \delta I + \epsilon_2 \leq 0 \end{cases} \tag{12}$$

Suppose we have incoherency because we believe $\gamma > 0$ (in the Hajivassiliou-Savignac study corresponding to binding FCs cause increasing chance of innovation I) and that $\delta > 0$ (firms who have high I i.e., innovate raise the chance the banks will refuse them a loan so high FC). So $\gamma \cdot \delta > 0$. This is Case 1 analyzed in subsection 7.3 as represented by Figure 2, and corresponding to the constraints on the unobservables:

$$(a^1, a^2)' < (\epsilon^1, \epsilon^2)' < (b^1, b^2)'$$

such that:

I	FC	a^1	b^1	a^2	b^2	Shading	Region
1	1	$-x_1\beta_1 - \gamma$	∞	$-x_2\beta_2 - \delta$	∞	horizontal	R1
1	0	$-x_1\beta_1$	∞	$-\infty$	$-x_2\beta_2 - \delta$	swne	R2
0	1	$-\infty$	$-x_1\beta_1 - \gamma$	$-x_2\beta_2$	∞	nwse	R3
0	0	$-\infty$	$-x_1\beta_1$	$-\infty$	$-x_2\beta_2$	vertical	R4

Now consider the transformed model with NFC **instead of** FC. This transformation still gives an overlapping region in the transformed variables, and hence corresponds to an incoherent model. To see this, proceed as follows:

In terms of the two latent variables I^* and $NFC^* = -FC^*$ and the observed binary indicators I and $NFC = 1 - FC$, and suppressing the observation index:

$$I = \begin{cases} 1 & if \quad I^* \equiv x_1\beta_1 + \gamma^N NFC + \epsilon_1 > 0 \\ 0 & if \quad I^* \equiv x_1\beta_1 + \gamma^N NFC + \epsilon_1 \leq 0 \end{cases} \tag{13}$$

$$NFC = \begin{cases} 1 & if \quad NFC^* \equiv x_2\beta_2^N + \delta^N I + \epsilon_2^N > 0 \\ 0 & if \quad NFC^* \equiv x_2\beta_2^N + \delta^N I + \epsilon_2^N \leq 0 \end{cases} \tag{14}$$

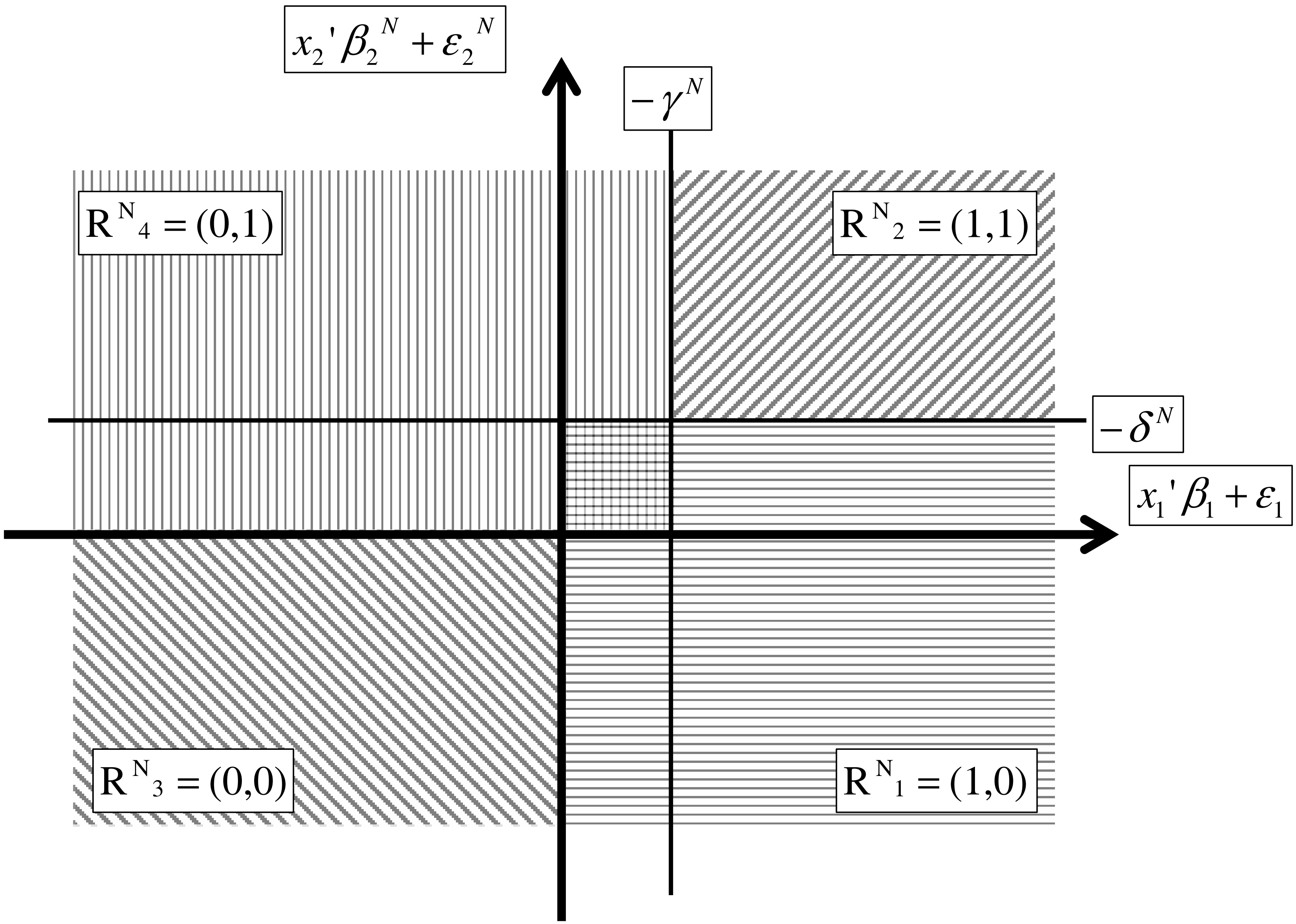
Given this transformation, we expect that $\gamma^N < 0$ (high NFC means not very binding constraints so cause dampening of I) and that $\delta^N < 0$ (firms who have high I i.e., innovate raise the chance the banks will refuse them a loan so low NFC). So $\gamma^N \cdot \delta^N > 0$.

For a typical i observation, the probability $Prob(y_{1i}, y_{2i}|X, \theta)$ is characterized by the constraints on the unobservables:

$$(a^1, a^2)' < (\epsilon_1, \epsilon_2^N)' < (b^1, b^2)'$$

through the configuration:

I	NFC	a^1	b^1	a^2	b^2	Shading	Region
1	0	$-x_1\beta_1$	∞	$-\infty$	$-x_2^N\beta_2 - \delta^N$	horizontal	R1
1	1	$-x_1\beta_1 - \gamma^N$	∞	$-x_2\beta_2^N - \delta^N$	∞	swne	R2
0	0	$-\infty$	$-x_1\beta_1$	$-\infty$	$-x_2^N\beta_2$	nwse	R3
0	1	$-\infty$	$-x_1\beta_1 - \gamma^N$	$-x_2^N\beta_2$	∞	vertical	R4



$x_2'\beta_2^N + \varepsilon_2^N$
$-\gamma^N$
$R^N_4 = (0,1)$
$R^N_2 = (1,1)$
$-\delta^N$
$x_1'\beta_1 + \varepsilon_1$
$R^N_3 = (0,0)$
$R^N_1 = (1,0)$

7.8 Efficient Estimation through Conditional Maximum Likelihood

The optimal parametric estimation approach for the models with empty region incoherency (Cases 3 and 4 above) will be *truncated conditional maximum likelihood*, employing the appropriate likelihood contributions that characterize correctly the necessary conditioning that ensures that the LDVs stay out of the empty region of incoherency. For example, assuming independence across observations $i = 1, \cdots, N$, the likelihood contribution in Case 3 will be:

$$l_i = Prob\,(\epsilon_1, \epsilon_2 : I = 1(I^* > 0)\ \&\ F = 1(F^* > 0))\,/\,(1 - Prob(-\gamma < \epsilon_1 + x_1\beta_1 < 0, 0 < \epsilon_2 + x_2\beta_2 < -\delta$$

while for Case 4:

$$l_i = Prob\,(\epsilon_1, \epsilon_2 : I = 1(I^* > 0)\ \&\ F = 1(F^* > 0))\,/\,(1 - Prob(0 < \epsilon_1 + x_1\beta_1 < -\gamma, \delta < \epsilon_2 + x_2\beta_2 < 0)$$

These likelihood contributions make it clear why approaches that ignore the coherency issue are inconsistent in general: the inconsistency would arise because the conditioning probability expressions in the denominator are functions of the underlying parameters and data, and hence affect critically the evaluation of the correct likelihood function.

It should be noted also that Cases 1 and 2 may be handled in an analogous fashion *provided it is assumed first that* the Data Generating Process that overcomes the overlapping-regions incoherency is one where $(\epsilon_{1i}, \epsilon_{2i})$ are drawn from an unrestricted bivariate normal distribution and then any draws falling into the overlap region are rejected. To find the correct likelihood contributions in these two cases, note that:

$$p^*_{11} + p^*_{10} + p^*_{01} + p^*_{00} = S > 1$$

where $S - 1 \equiv d$, the probability of the overlap region. In Case 1, the overlap occurs between regions $(1, 1)$ and $(0, 0)$, while for Case 2 between regions $(1, 0)$ and $(0, 1)$. Consequently, assuming an Accept/Reject DGP out of the overlap region, the likelihood contribution for Case 1 is:

$$l_i = \begin{cases} p_{11} = (p^*_{11} - d)/(2 - S) \\ p_{10} = p^*_{10}/(2 - S) \\ p_{01} = p^*_{01}/(2 - S) \\ p_{00} = (p^*_{00} - d)/(2 - S) \end{cases}$$

while for Case 2:

$$l_i = \begin{cases} p_{11} = p^*_{11}/(2 - S) \\ p_{10} = (p^*_{10} - d)/(2 - S) \\ p_{01} = (p^*_{01} - d)/(2 - S) \\ p_{00} = p^*_{00}/(2 - S) \end{cases}$$

8 Extensions to Bivariate Multinomial Ordered Probit Cases

We now discuss how to extend our analysis to the case of two simultaneous (bivariate) *ordered probit equations with multiple regions.*[5] Suppose we have a model given by

$$\begin{aligned} y_1^* &= \beta_1' x_1 + \delta_{y_2} + \epsilon_1 \\ y_2^* &= \beta_2' x_2 + \delta_{y_1} + \epsilon_2 \\ y_1 &= I_1(y_1^*) \\ y_2 &= I_2(y_2^*) \end{aligned}$$

where we define

$$\begin{aligned} I_1(y_1^*) &\in \{1, 2, ..., n_1\} \\ I_2(y_2^*) &\in \{1, 2, ..., n_2\} \\ I_1(y_1^*) &= \max\{i | y_1^* < s_{1i}\} \\ I_2(y_2^*) &= \max\{i | y_2^* < s_{2i}\} \end{aligned}$$

The sets $\{s_{1i}\}$ and $\{s_{2i}\}$ are $n_1 - 1$ and $n_2 - 1$ increasing transition values, and $s_{1n_1i} = s_{2n_2} = \infty$, so that all very large values get mapped to the highest category. Then

$$\begin{aligned} \delta_{y_1} &\in \{\delta_{11}, \delta_{12}, ..., \delta_{1n_2}\} \\ \delta_{y_2} &\in \{\delta_{21}, \delta_{22}, ..., \delta_{1n_1}\} \end{aligned}$$

are interaction terms that take one of n_2 and n_1 discrete values depending on y_2 and y_1 respectively. The error terms, ϵ_1 and ϵ_2, are assumed to be normally distributed conditional on lying outside of incoherent regions – that is, regions in which there is not a single, unambiguous pair y_1 and y_2 that corresponds to them.

The bivariate binary probit is a subset of this case, with $n_1 = n_2 = 2$, $s_{11} = s_{21} = 0$, $\delta_{11} = \gamma$, $\delta_{21} = \delta$, using our usual notation.

As an illustration, consider the following figure. Here, $n_1 = n_2 = 3$, and all the $\delta_{2i} = 0$ so that y_1 does not affect y_2. As in the binary probit, the effect of the interaction terms are to shift the boundaries in ϵ's domain that map to particular outcomes for y_1 and y_2. For example, the shaded area corresponds to $y_1 = 2$, $y_2 = 2$.

In this picture, there are no incoherent regions, since the system is triangular. To proceed to determine the regions of incoherency in the non-triangular case, for

[5] The research assistance of Ryan Giordano has been especially helpful for this section.

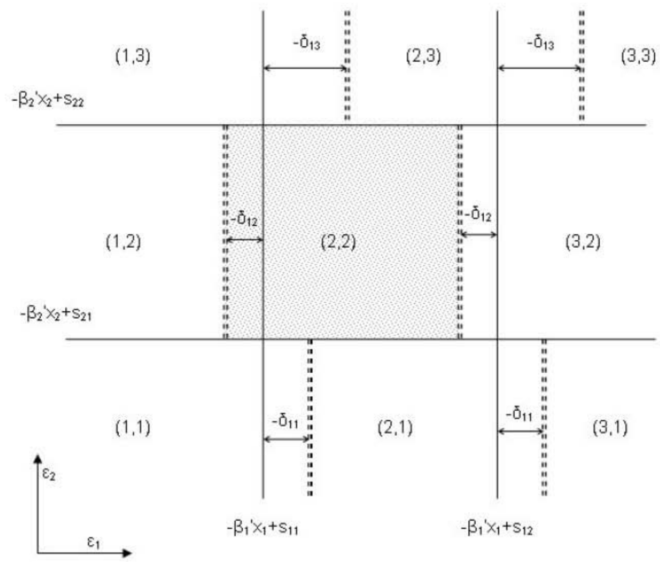


Figure 7: Coherency of Joint Ordered Response Models

simplicity we will assume that the interaction terms are small enough so that $s_{1(i+1)} - s_{1i} > \delta_{2(j+1)} - \delta_{1j}$ for all i and j, and that a similar condition holds for the transition values and interaction terms of y_2. That is to say, the interaction terms are small relative to the threshold values. If this is not the case, counting overlaps becomes more complicated.

If this condition holds, then each intersection of threshold values becomes analogous to the binary bivariate probit case, except that neighbouring interaction terms now determine whether there is an empty region or overlap. For example, consider the following situation. Here,

$$
\begin{aligned}
n_1 &= 3 \\
n_2 &= 2 \\
\delta_{11} &< 0 \\
\delta_{12} &> 0 \\
\delta_{21} &> 0 \\
\delta_{22} &< 0 \\
\delta_{23} &> 0
\end{aligned}
$$

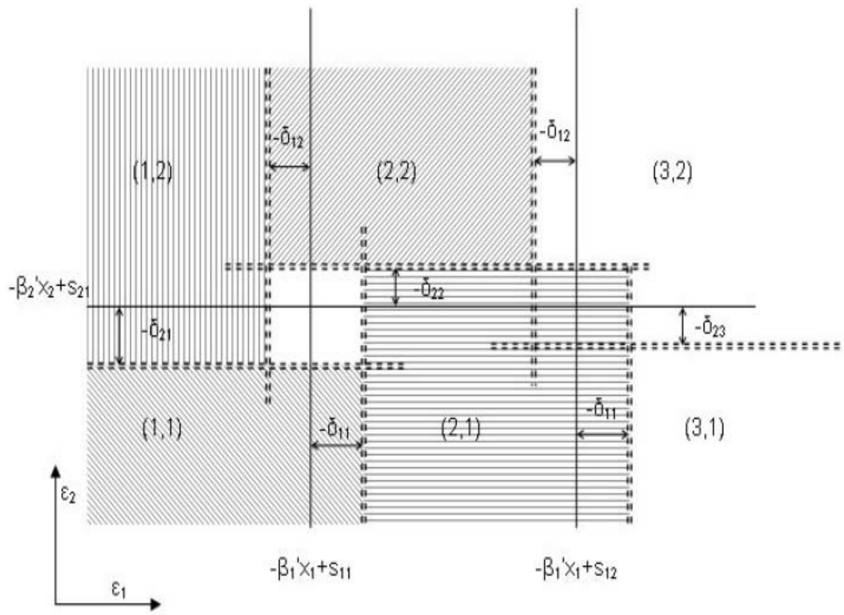


Figure 8: Coherency of Joint Ordered Response Models II

When we draw in the shaded regions for the first corner, it is evident that there is an empty region. This is because $\delta_{12} - \delta_{11} > 0$ has the same a different sign from $\delta_{22} - \delta_{21} < 0$. Indeed, if we were to take

$$z_1 = -\beta_1' x_1 - \delta_{11} z_2 \qquad = -\beta_2' x_2 - \delta_{21} \gamma_1 = \delta_{12} - \delta_{11}$$
$$\gamma_2 = \delta_{22} - \delta_{21}$$

then the situation would be identical to the bivariate binary probit with the role of the exogenous variables played by z_1 and z_2, and the relevant interaction terms being γ_1 and γ_2. Here, for similar reasons, we will see an overlap at the second intersection:

Having observed this, it is easy to define a procedure to calculate the overall likelihood with coherency incorporated. For a particular outcome, one first calculates its incoherent probability (the probability of (ϵ_1, ϵ_2) landing in its bounding box). Then one checks each of its corners for overlaps with neighbouring regions, subtracting the probability of any overlap regions. Finally, one needs to divide by the total probability of a coherent draw, which equals one minus the sum of the probabilities of all the empty and overlapping regions.

We have investigated the theoretical probabilities for the model made coherent in this fashion, and compared them to the actual frequency probabilities from the

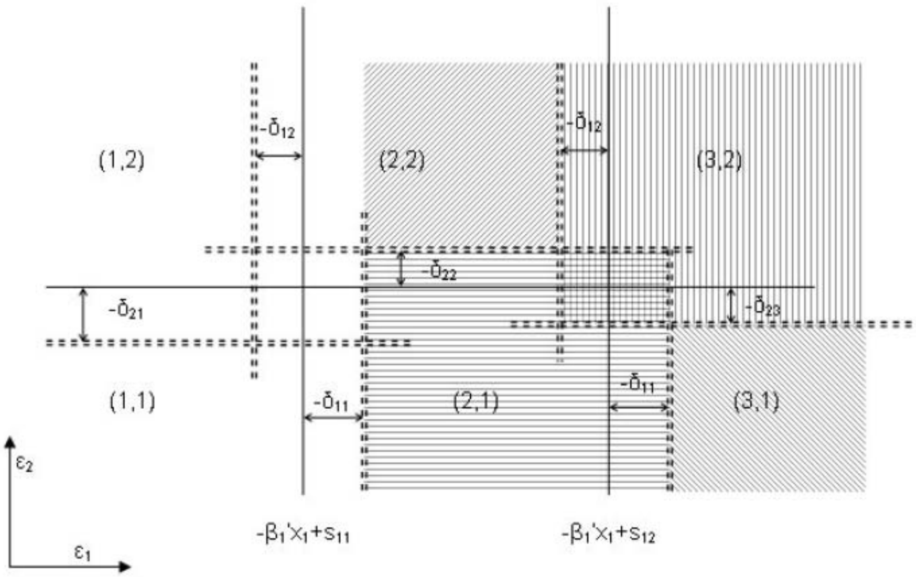


Figure 9: coherency of joint ordered Response Models III

postulated accept-reject DGP described in the first paragraph of this section. The two sets of probabilities generated in these investigations matched one another very satisfactorily.

8.1 Establishing the Coherency of Panel LDV Models with Intertemporal Endogeneities using DGP Approach

Extending the analysis to a panel data set, Hajivassiliou (2007)[6] explains how the probability of a pair (S_{it}, E_{it}) in subsection 4.1, a pair (C_{it}, B_{it}) in subsection 4.2, and a pair (I_{it}, F_{it}) in subsection 4.3, can be represented in terms of the linear inequality:

$$(a_1, a_2)' < (\epsilon_1, \epsilon_2)' < (b_1, b_2)'$$

where the error vector has a flexible autocorrelation structure. For example, one-factor random effect assumptions will imply an equicorrelated block structure on Σ_ϵ, while our most general assumption of one-factor random effects *combined with* an AR(1) process for each error implies that Σ_ϵ combines equicorrelated and Toeplitz-matrix features. Consequently, the approach incorporates fully (a) the contemporaneous correlations in ϵ_{it}, (b) the one-factor plus AR(1) serial correlations in ϵ_i, and (c) the dependency of S_{it} on E_{it} and vice versa. The coherency issue expands naturally to the panel sequence of data, by thinking of each (correlated) time-period for a given individual i as a **distinct probit equation** and then dealing with the independent cross-section of equations across individuals. Details of the analysis can be found in Hajivassiliou (op.cit.).

Our hypothetical DGP method presented in Subsection 6.3 for establishing coherency is now applied to the canonical panel data Probit model with state-dependence, first analyzed by Heckman (op.cit.). Let us begin with the simplified case of the initial condition being exogenous:

$$y_{iT} = \mathbf{1}(\lambda y_{i,T-1} + x_{iT}\beta + \epsilon_{iT} > 0) \tag{15}$$

$$y_{i,T-1} = \mathbf{1}(\lambda y_{i,T-2} + x_{i,T-1}\beta + \epsilon_{i,T-1} > 0) \tag{16}$$

$$\vdots \tag{17}$$

$$y_{i2} = \mathbf{1}(\lambda y_{i1} + x_{i2}\beta + \epsilon_{i2} > 0) \tag{18}$$

$$y_{i1} = \textbf{exogenous} \tag{19}$$

Let $\Sigma \equiv VCov(\epsilon_{iT}, \cdots, \epsilon_{i1}, u_{i1})$. Suppose first the ϵ_{it} has the one-factor (equicorrelated) error components structure $\epsilon_{it} = \alpha_i + \nu_{it}$. Conditional on α_i, these $T-1$ equations are independent (since they only depend on the i.i.d. ν_{it}s. Hence draw an α_i and an independent ν_{i2}. Then use the exogenous y_{i1} outcome to generate y_{i2}. This completes equation 18 which allows to move recursively to generating y_{i3}, then y_{i4}, etc. until y_{iT} is generated. This establishes the coherency of the model.

Now consider the more general case when y_{i1} cannot be assumed as exogenous. We then supplement the system with an initial condition equation:

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + u_{i1} > 0$$

The following remarks are in order: First note that equation 1 is a generalization of the Barghava and Sargan (1982)[1] approach. Second, one-factor random effect assumptions will imply an equicorrelated block structure on the top left $T-1 \times T-1$ block of Σ, while more general assumptions of one-factor random effects *combined with* an AR(1) or ARMA(p,q) processes for each ϵ error implies that Σ combines equicorrelated and Toeplitz-matrix parts. The last row and last column of Σ giving the variance of u_{1i} and its covariances with all ϵ_{it} allow the flexibility stipulated by Heckman (1981b)[12]. Define the Cholesky lower triangular times upper triangular factorization of $\Sigma = CC'$. Given the assumed normality, the error vector can be written:

$$(\epsilon_i', u_{1i})' = C\nu_i \qquad \nu_i \sim N(0_T, I_T) \tag{20}$$

Dropping the i index:

$$y_T = \mathbf{1}(\lambda y_{T-1} + x_T\beta + c_{T1}\nu_1 + c_{T2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} + c_{TT}\nu_T > 0) \tag{21}$$

$$y_{T-1} = \mathbf{1}(\lambda y_{T-2} + x_{T-1}\beta + c_{T-1,1}\nu_1 + c_{T-1,2}\nu_2 + \cdots + c_{T,T-1}\nu_{T-1} > 0) \tag{22}$$

$$\vdots \tag{23}$$

$$y_2 = \mathbf{1}(\lambda y_1 + x_2\beta + c_{22}\nu_2 + c_{21}\nu_1 \tag{24}$$

$$y_{i1} = \mathbf{1}(x_{i1}\xi_1 + \cdots + x_{iT}\xi_T + c_{11}\nu_{i1} > 0) \tag{25}$$

This recursive representation establishes the coherency of the model: given a random draw of $\nu_{i1}, \cdots, \nu_{iT}$, an unambiguous DGP rule can be defined to establish sequentially $y_{i1} \rightarrow y_{i2} \rightarrow \cdots y_{i,T-1} \rightarrow y_{iT}$.

9 Monte-Carlo Experiments: Design

As we showed in the previous section, we obtain a coherent non-recursive model with interaction dummy included on each side, provided we believe the feedback terms have opposite signs on the two sides. We also showed how to handle the cases of the feedback terms having the same sign through the additional assumption of the accept-reject DGP.

The experiments were designed to illustrate the importance of coherency on the following *nine* estimation approaches:

(a) likelihood estimation that incorrectly forces the old coherency condition to hold, i.e., assuming recursivity when in fact both feedback terms are present (estimators `E-TRWN`=assuming $\delta = 0$ and `E-TRNW`=assuming $\gamma = 0$);

(b) unrestricted likelihood estimation, which ignores the resulting incoherency due to the empty or overlap region(s) (estimator `E-INCO`);

(c) restricted likelihood estimation conditioning on the data lying outside the empty region(s) of incoherency (estimators `E-SQPM`=assuming $(\gamma \geq 0, \delta \leq 0)$ and `E-SQMP`=assuming $(\gamma \leq 0, \delta \geq 0)$);

(d) restricted likelihood estimation conditioning on the data lying outside the overlap region(s) of incoherency (estimators `E-SQPP`=assuming $(\gamma \geq 0, \delta \geq 0)$ and `E-SQMM`=assuming $(\gamma \leq 0, \delta \leq 0)$).

(e) LPOLS: (linear probability) ordinary least squares estimation of each binary probit equation ignoring the possible endogeneity of the interaction terms; and LP2SLS: applying two-stage least squares recognizing that the two interaction terms on the RHS of each probit equation can be endogenous.

We generate six "true" models:

- `DGP-TRWN` $(\delta = 0)$
- `DGP-TRNW` $(\gamma = 0)$
- `DGP-SQPM` $(\gamma \geq 0, \delta \leq 0)$
- `DGP-SQMP` $(\gamma \leq 0, \delta \geq 0)$
- `DGP-SQPP` $(\gamma \geq 0, \delta \geq 0)$ and
- `DGP-SQMM` $(\gamma \leq 0, \delta \leq 0)$,

and in each case, calculate the nine estimators `E-TRWN`, `E-TRNW`, `E-INCO`, `E-SQPM`, `E-SQMP`, `E-SQPP`, `E-SQMM, LPOLS,` and `LP2SLS.`

The generating equations are:

$$ystar1 = x1[nobs, kx1] * beta1 + gamma * y2 + eps1, \quad y1 = 1(ystar1 > 0)$$

$$ystar2 = x2[nobs, kx2] * beta2 + delta * y1 + eps2, \quad y2 = 1(ystar2 > 0)$$

9.1 γ unrestricted, $\delta = 0$

$$ystar1 = x1[nobs, kx1] * beta1 + gamma * y2 + eps1, \quad y1 = 1(ystar1 > 0)$$

$$ystar2 = x2[nobs, kx2] * beta2 + eps2, \quad y2 = 1(ystar2 > 0)$$

Given the recursivity of the $\gamma \cdot \delta = 0$ restriction in this case, $ystar2$ is generated first, which gives $y2$. This is then plugged into the RHS of the $ystar1$ equation thus allowing $ystar1$ and $y1$ to be obtained.

9.2 $\gamma \geq 0, \delta \leq 0$

$$0 \leq eps1 + x1 * b1 \leq gamma, -delta \leq eps2 + x2 * beta2 \leq 0 \tag{26}$$

Accept-reject methods are used to generate the data so that these restrictions are satisfied.

Analogous Accept/Reject DGP for the $\gamma \geq 0, \delta \geq 0$ case. Also see appendix 1 for an exact algorithm for generating draws from truncated normal distributions restricted to lie on region (26).

9.3 $\gamma \leq 0, \delta \geq 0$

$$-gamma \leq eps1 + x1 * b1 \leq 0, 0 \leq eps2 + x2 * beta2 \leq -delta$$

Accept-reject methods are used to generate the data so that these restrictions are satisfied.

Analogous Accept/Reject DGP for the $\gamma \leq 0, \delta \leq 0$ case.

10 Monte Carlo Experiments: Findings

We performed 24 Monte-Carlo experiments, indexed by MCxyz as follows:

	δ	γ
$x = 1$	0	0
$x = 2$	0.8	0
$x = 3$	0.8	1
$x = 4$	0.8	-1

	$\rho_{\epsilon_1,\epsilon_2}$
$y = 1$	0.3
$y = 2$	-0.3

	x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	x_{23}
$z = 1$	$const$	$\chi^2(1)$	$Bernoulli(0.7)$	$const$	x_{12}	$DoubleExponentialSS$
$z = 2$	$const$	$\chi^2(1)$	$Bernoulli(0.9)$	$const$	x_{12}	$DoubleExponentialSS$
$z = 3$	$const$	$\chi^2(1)$	$Bernoulli(0.7)$	$const$	x_{12}	$DoubleExponentialLS$
$z = 4$	$const$	$\chi^2(1)$	$Bernoulli(0.9)$	$const$	x_{12}	$DoubleExponentialLS$

where *DoubleExponential* stands for a Double Exponential distribution with mean 0 with asymmetric two sides, *SS* for "small skewness" and *LS* with "large skewness."

All trials used 2000 observations and 200 Monte-Carlo trials were averaged in each case. In all experiments, the true beta parameters were set at: $\beta_1 = \begin{pmatrix} 0.8, & -0.5, & -0.3 \end{pmatrix}'$ and $\beta_2 = \begin{pmatrix} -0.3, & 0.7, & -0.4 \end{pmatrix}'$. In Appendix 2 below, we give the complete listing of the regime probabilities in all 32 experiments we carried out.

The full tables presenting the detailed Monte-Carlo results in terms of various estimation criteria (root-mean-squared error, absolute bias, absolute median bias, variance, interquartile range, and nine-decile range) can be obtained from the author upon request. Here we summarize the results in terms of three dimensional bar charts in four-part Figures 10-16 below. A given chart graphs the performance of each estimation algorithm in terms of a given estimation criterion (e.g., RMSE etc.) with the best performing algorithm normalized to 100. The other methods are then given as a fraction of that best. For example, if method A is the best with RMSE=25, and methods B and C have RMSE equal to 75 and 125 respectively, method A will be reported as 100, B as 0.333 (one third as good since 3 times as high RMSE), and C as 0.20 (one fifth as good since 5 times as high RMSE).

We first give a very drastic summary of the main findings:

- The Conditional Truncated MLE proposed in this paper performs very satisfactorily, being the only consistent estimator for the reverse feedback cases, and only small sacrifices in terms of efficiency in the recursive DGPs when it is not strictly necessary.

- The linear probability estimators, LPOLS and LOP2SLS, perform very badly in all cases with endogenous interaction terms, thus suggesting that the inherent non-linearities of the bivariate probits cannot be safely ignored.

- Conditional Truncated MLE also works well for the overlap region incoherency cases.

- Unrestricted likelihood estimation ignoring the resulting incoherency due to the empty or overlap region(s) (estimator `E-INCO`) is by far the worst performing estimator, dominated even by equation by equation univariate estimators which ignore the other side of the model.

  More analytically:

- The four-part Figures 10 present the *overall* RMSE results with each method's performance averaged across all estimated parameters. The CMLE estimator dominates all other methods in impressive fashion when the true DGP possesses the opposite-signs restriction $\gamma \cdot \delta \leq 0$. It also performs very satisfactorily in case the true model is recursive, achieving almost as good a performance as the ideal recursive estimator for that case. Even in the case of no interaction terms being present in the true DGP ($\gamma \cdot \delta = 0$), the CMLE estimator loses out in terms of RMSE only because of the higher estimation variance in view of not imposing two true restrictions.

- The four-part Figures 11 report relative RMSE performance for the δ interaction parameter, whereas Figures 12, 13, and 14 give the results for the β_{11}, β_{22}, and ρ respectively. CMLE also impresses in these sets of results in a similar ranking to the previous point.

- The four-part Figures 14 present the *overall* results in terms of *absolute bias* instead of RMSE, whereas Figures 15 give the *overall* results in terms of absolute *median* bias. The first set establishes that the CMLE estimator heads and shoulders above all the alternatives in terms of bias, and whenever it is less clearly the preferred estimator, this only caused by higher estimation variance.

- Figures 15 allow one to draw conclusions about the extend of non-symmetry of the distributions of the alternative estimators. No dramatic changes in the rankings of estimator performance are apparent in this regard.

It may be noted that the dismal performance of the two estimators based on the Linear Probability approximation would have been alleviated had the average partial probability derivatives been calculated instead of the latent variable coefficients. This is because the LP estimators by construction a constant probability derivative with respect to an explanatory variable, irrespective of the observation values. In our view, such calculations would not be especially interesting since in most empirical LDV studies, investigators wish to allow for such probability derivatives to vary over the range of observations.

11 Conclusions

The paper discussed the major identification issue of *coherency conditions* in LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables. The econometric framework of LDV models with simultaneity was presented and the identification issue of *coherency* in such LDV models with endogeneity and flexible temporal and contemporaneous correlations in the unobservables was analyzed.

Conditions for coherency as presented in the existing literature were reviewed and shown to be rather esoteric. Two novel methods for establishing coherency conditions were presented, one based on a graphical characterization, the second through hypothetical Monte-Carlo DGP. The novel approaches have intuitive interpretations and are easy to implement and generalize. The constructive consequence of the new approaches is that they indicate how to achieve coherency in models traditionally classified as incoherent through the use of prior sign restrictions on model parameters. This allowed us to develop estimation strategies based on Conditional MLE for simultaneous LDV models without imposing recursivity. Thus one can obtain for the first time estimates of direct as well as reverse interaction effects in simultaneous LDV models, unlike in the existing literature where recursivity had to be assumed. Econometric applications were used to illustrate the methods in practice and extensions are given to simultaneous ordered probit models with multiple regions.

The proposed Conditional MLE methodology was evaluated through an extensive set of Monte-Carlo experiments. The experiments allowed us also to study the consequences of employing estimators that make overly restrictive coherency assumptions about the DGP. The findings confirmed very substantive improvements in terms of estimation Mean-Squared-Error by employing the CMLE developed in this paper. They also showed that estimators based on the Linear Probability approximation perform poorly in this context.

Our CMLE approach allows for the first time to obtain estimates of the reverse as well as direct interaction terms in LDV models with simultaneity.

12 Appendix 1: Generating standard Normal variates truncated to lie *outside* $[\underline{\lambda}, \bar{\lambda}]$

We present here a method for generating truncated normal variates to ensure the coherency of the non-recursive model under prior sign restrictions:

Let $z \sim N(0,1)$ and define $\tau \sim z|\{z \notin [\underline{\lambda}, \bar{\lambda}]\}$ Then $cdf(z) : F(z) = \Phi(z)$ and

$$cdf(\tau) : F(\tau) = \begin{cases} \frac{\Phi(z)}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } z < \underline{\lambda}, \\ \frac{\Phi(\underline{\lambda})}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } \underline{\lambda} < z \leq \bar{\lambda}, \\ \frac{\Phi(z)-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})}{1-\Phi(\bar{\lambda})+\Phi(\underline{\lambda})} & \text{if } z > \bar{\lambda}. \end{cases}$$

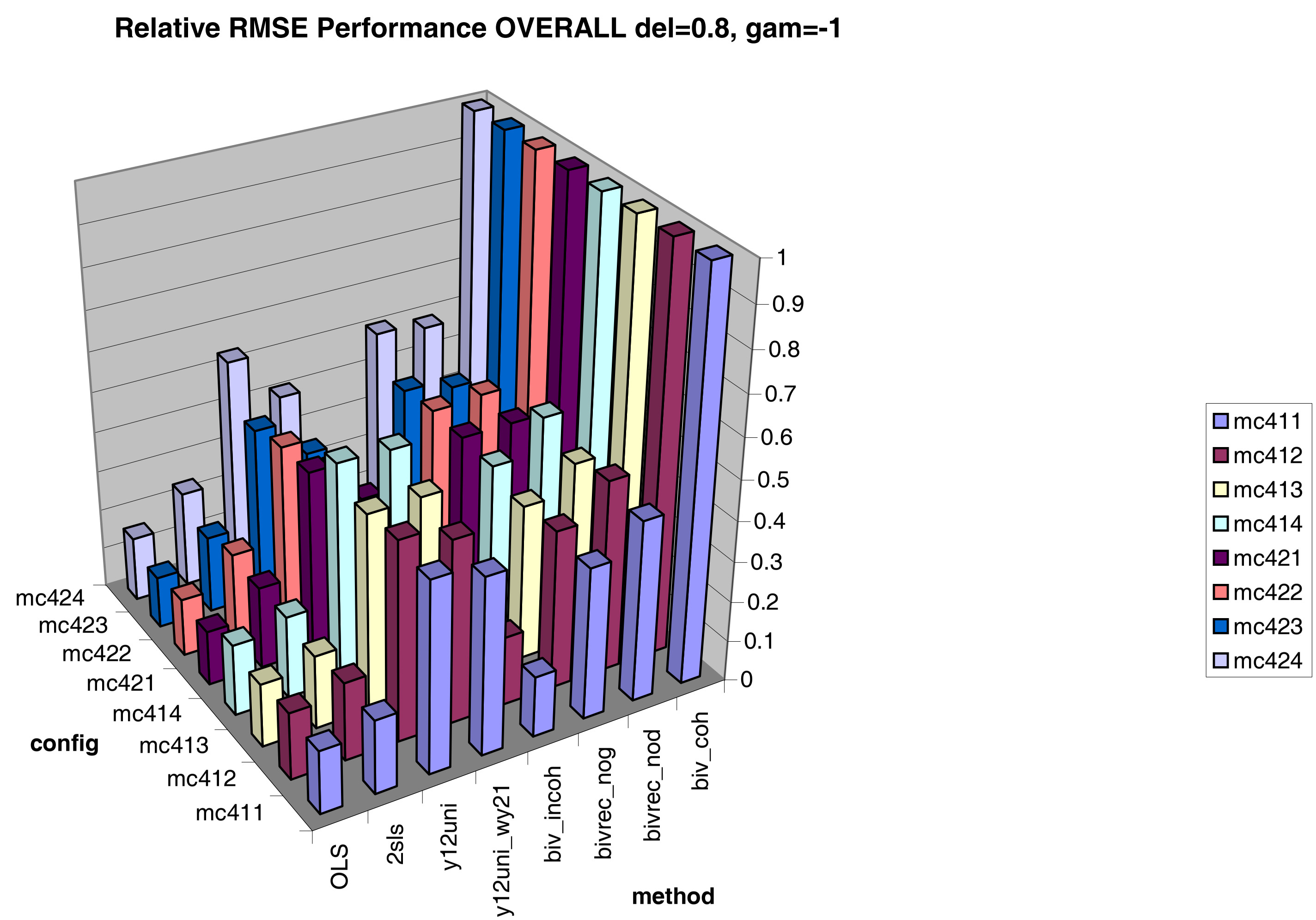
The procedure is exact for a univariate z truncated on $\{z \notin [\underline{\lambda}, \bar{\lambda}]\}$, but it will not work for higher dimensions. For DGPs with higher dimensions, accept-reject methods are preferable, though others exist (e.g., Gibbs resampling — see Hajivassiliou and McFadden (1998) for an explanation.

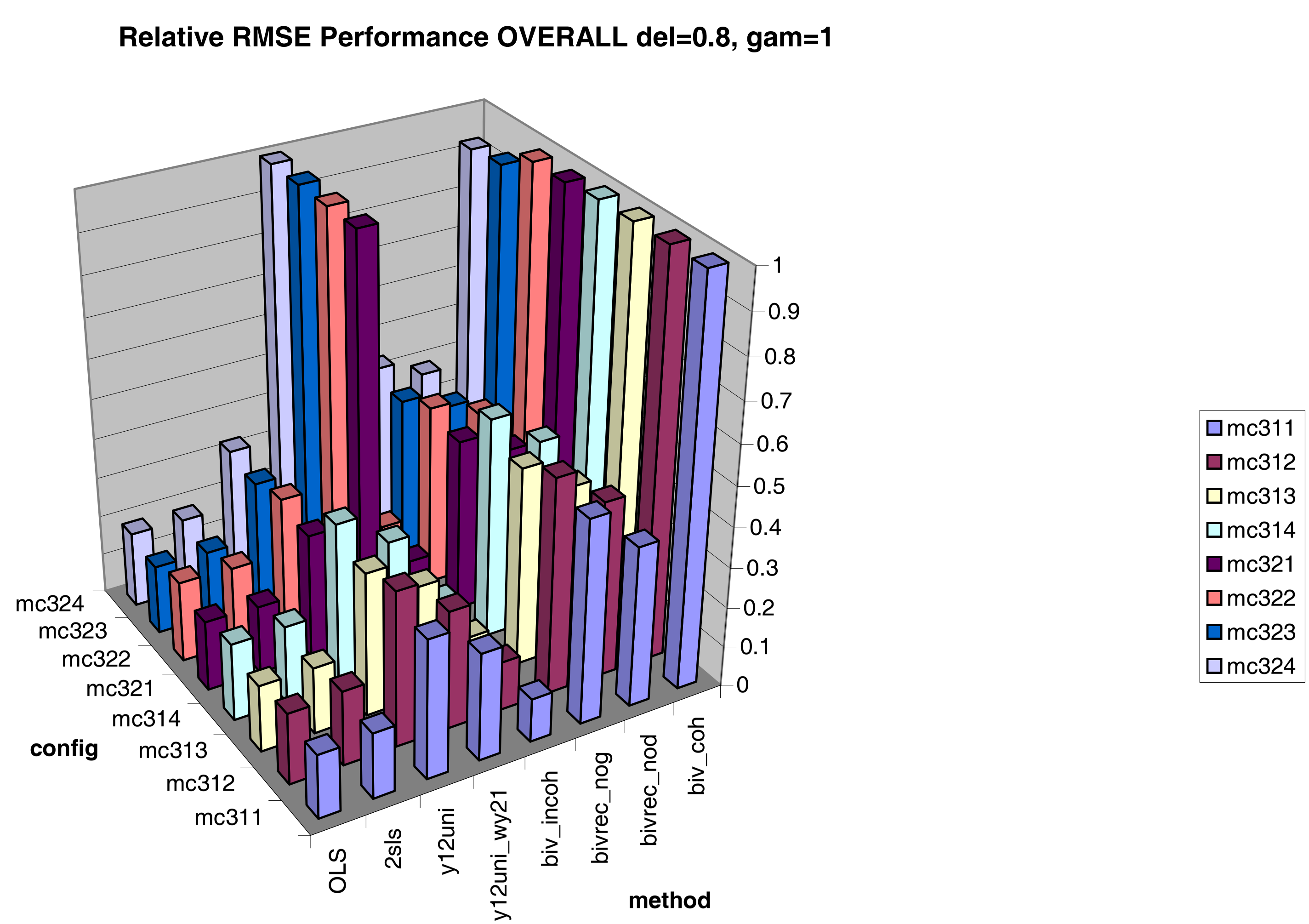
13 Appendix 2: Regime Probabilities in Monte-Carlo Experiments

	$Y_2 = 1$	$Y_2 = 0$	
$Y_1 = 1$	p_{11}	p_{10}	$p_{1\cdot}$
$Y_1 = 1$	p_{01}	p_{00}	$p_{0\cdot}$
	$p_{\cdot 1}$	$p_{\cdot 0}$	

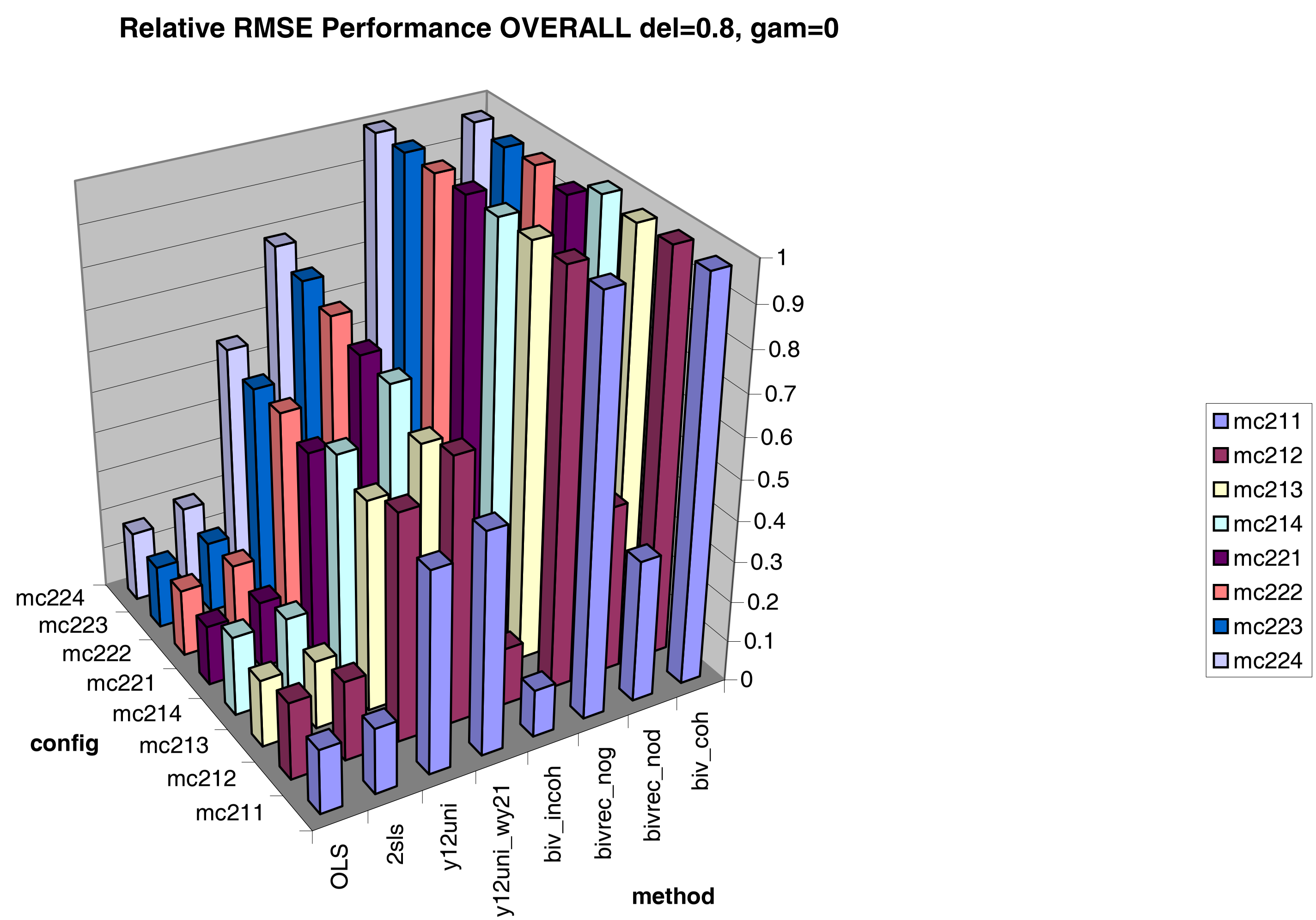
Experiment	p_{11}	p_{10}	p_{01}	p_{00}	$p_{0\cdot}$	$p_{1\cdot}$	$p_{\cdot 1}$	$p_{\cdot 0}$
mc111	0.2812	0.2716	0.1759	0.2711	0.4470	0.5529	0.4572	0.5427
mc112	0.2736	0.2589	0.1845	0.2829	0.4674	0.5325	0.4581	0.5418
mc113	0.2696	0.2840	0.1748	0.2714	0.4463	0.5536	0.4445	0.5554
mc114	0.2598	0.2728	0.1844	0.2829	0.4673	0.5326	0.4442	0.5557
mc121	0.2262	0.3273	0.2316	0.2147	0.4464	0.5535	0.4579	0.5421
mc122	0.2175	0.3162	0.2400	0.2261	0.4662	0.5337	0.4576	0.5424
mc123	0.2219	0.330	0.2229	0.2242	0.4472	0.5527	0.4449	0.5551
mc124	0.2130	0.321	0.2306	0.2350	0.4657	0.5342	0.4437	0.5563
mc211	0.4054	0.148	0.1751	0.2706	0.4458	0.5541	0.5806	0.4419
mc212	0.3920	0.141	0.1852	0.2816	0.4669	0.5330	0.5772	0.4227
mc213	0.3772	0.176	0.1757	0.2709	0.4466	0.5533	0.5530	0.4470
mc214	0.3661	0.167	0.1835	0.2829	0.4665	0.5334	0.5497	0.4503
mc221	0.3515	0.201	0.2317	0.2153	0.4471	0.5528	0.5833	0.4167
mc222	0.3381	0.196	0.2405	0.2252	0.4658	0.5341	0.5786	0.4213
mc223	0.3275	0.225	0.2219	0.2251	0.4471	0.5528	0.5495	0.4505
mc224	0.3141	0.218	0.2327	0.2349	0.4676	0.5323	0.5468	0.4532
mc311	0.5523	0.157	0.0652	0.22	0.2899	0.7100	0.6175	0.3824
mc312	0.5441	0.149	0.0696	0.2368	0.3064	0.6935	0.6138	0.3862
mc313	0.5163	0.185	0.0663	0.2319	0.2983	0.7016	0.5826	0.4173
mc314	0.5080	0.177	0.0712	0.2429	0.3142	0.6857	0.5793	0.4207
mc321	0.5155	0.218	0.0998	0.1659	0.2658	0.7341	0.6154	0.3845
mc322	0.5070	0.211	0.1058	0.1758	0.2816	0.7183	0.6128	0.3871
mc323	0.4818	0.240	0.0942	0.1836	0.2778	0.7221	0.5761	0.4238
mc324	0.4726	0.233	0.1007	0.1931	0.2939	0.7060	0.5734	0.4266
mc411	0.1903	0.163	0.3520	0.2937	0.6457	0.3542	0.5423	0.4576
mc412	0.1773	0.155	0.3607	0.3066	0.6674	0.3325	0.5381	0.4619
mc413	0.1773	0.190	0.3403	0.2919	0.6322	0.3677	0.5176	0.4823
mc414	0.1638	0.181	0.3506	0.3036	0.6543	0.3456	0.5145	0.4854
mc421	0.1430	0.223	0.4017	0.2317	0.6334	0.3665	0.5447	0.4552
mc422	0.1308	0.216	0.4096	0.2432	0.6529	0.3470	0.5405	0.4595
mc423	0.1331	0.244	0.3829	0.2392	0.6221	0.3778	0.5161	0.4839
mc424	0.1201	0.237	0.3909	0.2515	0.6425	0.3574	0.5111	0.4889

Figure 10

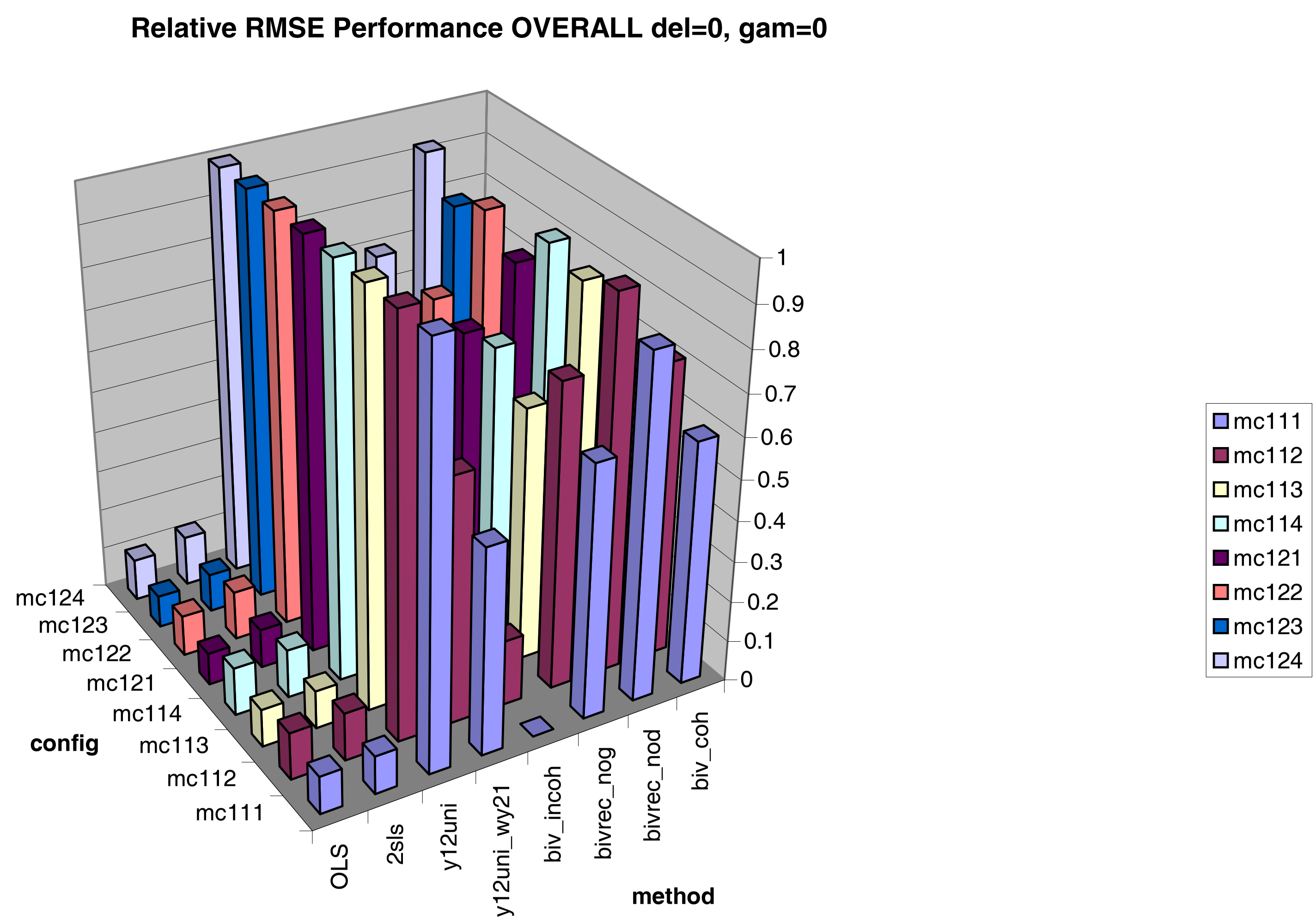

Relative RMSE Performance OVERALL del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424



Relative RMSE Performance OVERALL del=0.8, gam=1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc324
mc323
mc322
mc321
mc314
mc313
mc312
mc311
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc311
mc312
mc313
mc314
mc321
mc322
mc323
mc324

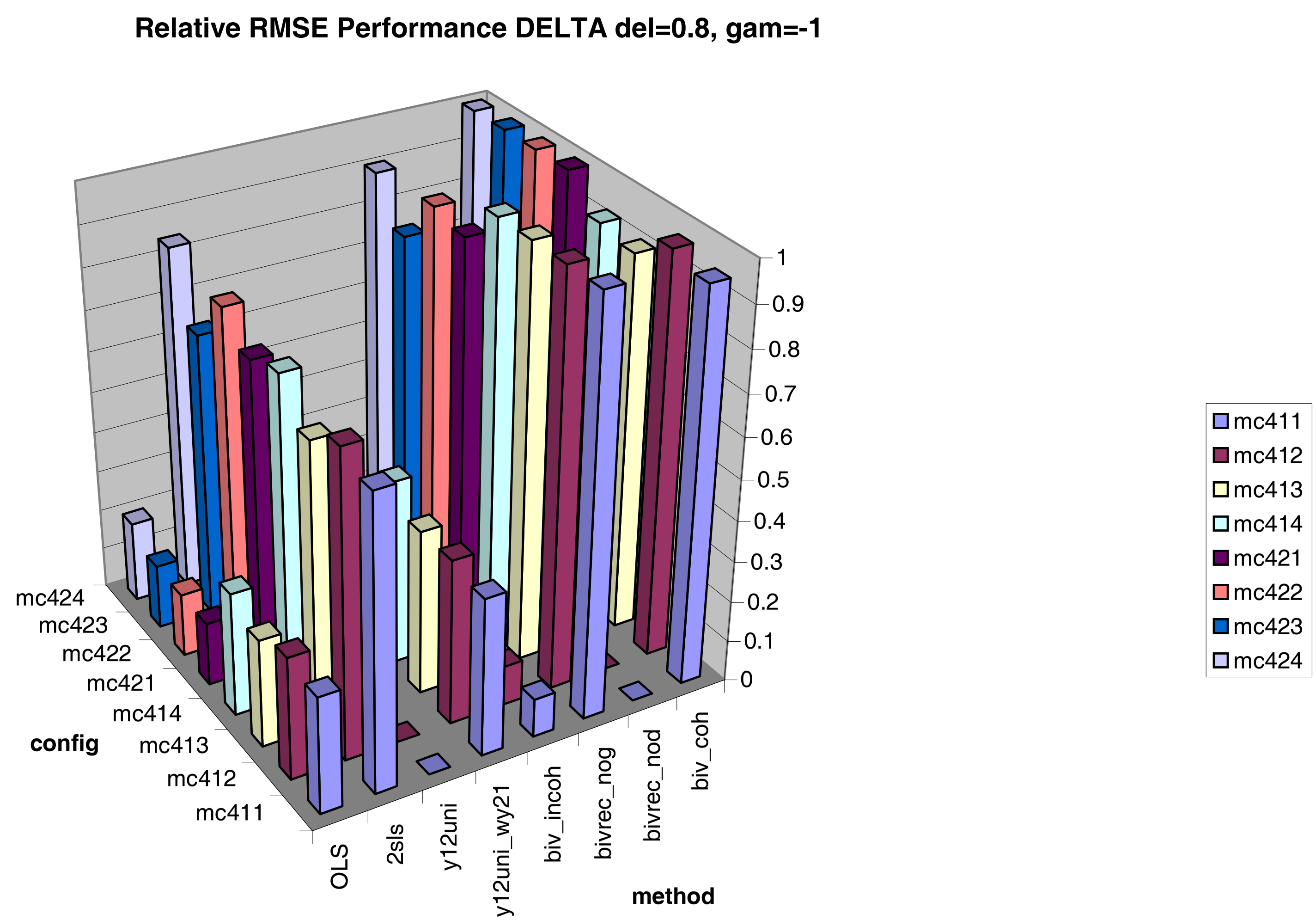


Relative RMSE Performance OVERALL del=0.8, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc224
mc223
mc222
mc221
mc214
mc213
mc212
mc211
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc211
mc212
mc213
mc214
mc221
mc222
mc223
mc224



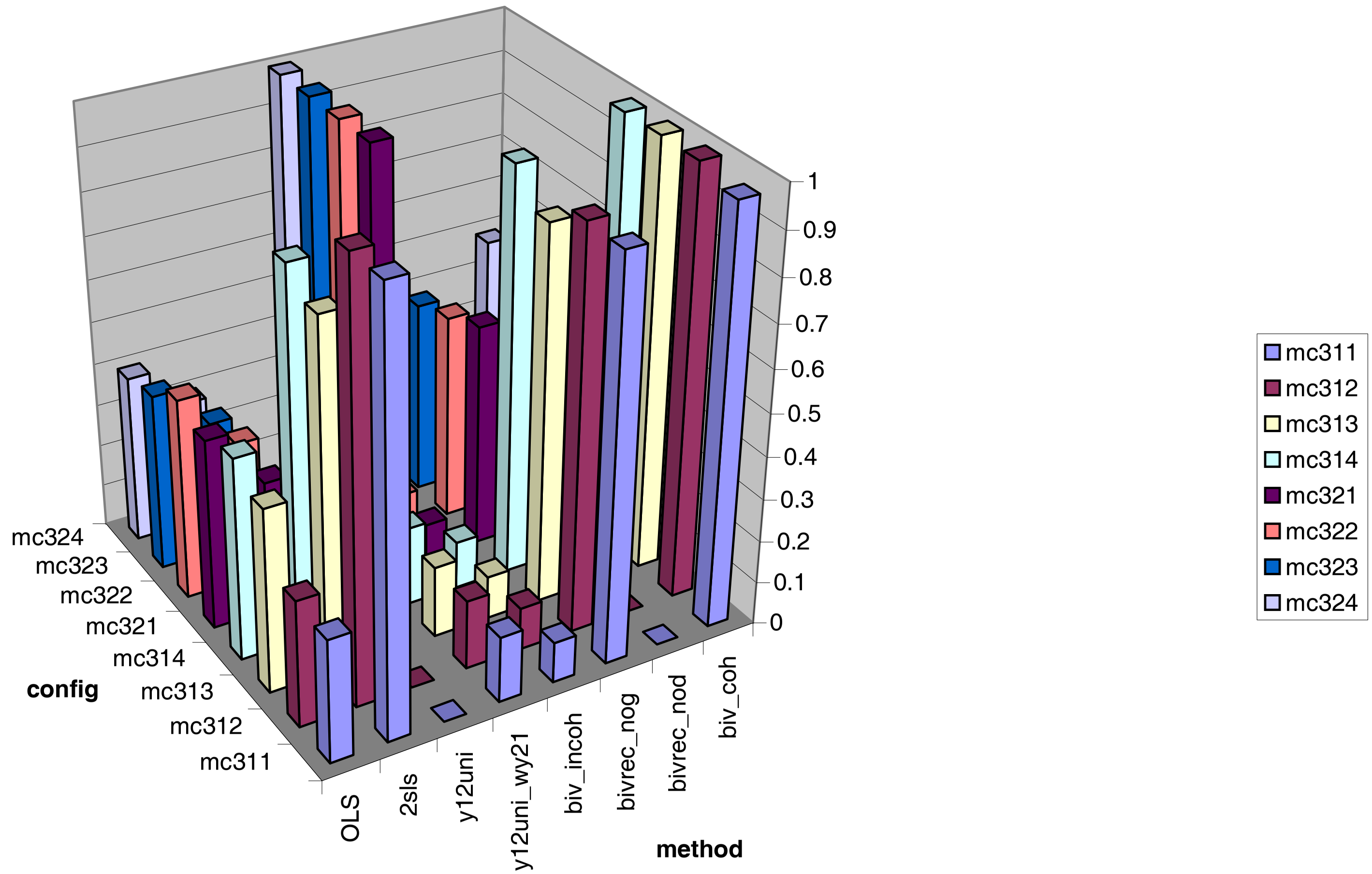
Relative RMSE Performance OVERALL del=0, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc124
mc123
mc122
mc121
mc114
mc113
mc112
mc111
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc111
mc112
mc113
mc114
mc121
mc122
mc123
mc124

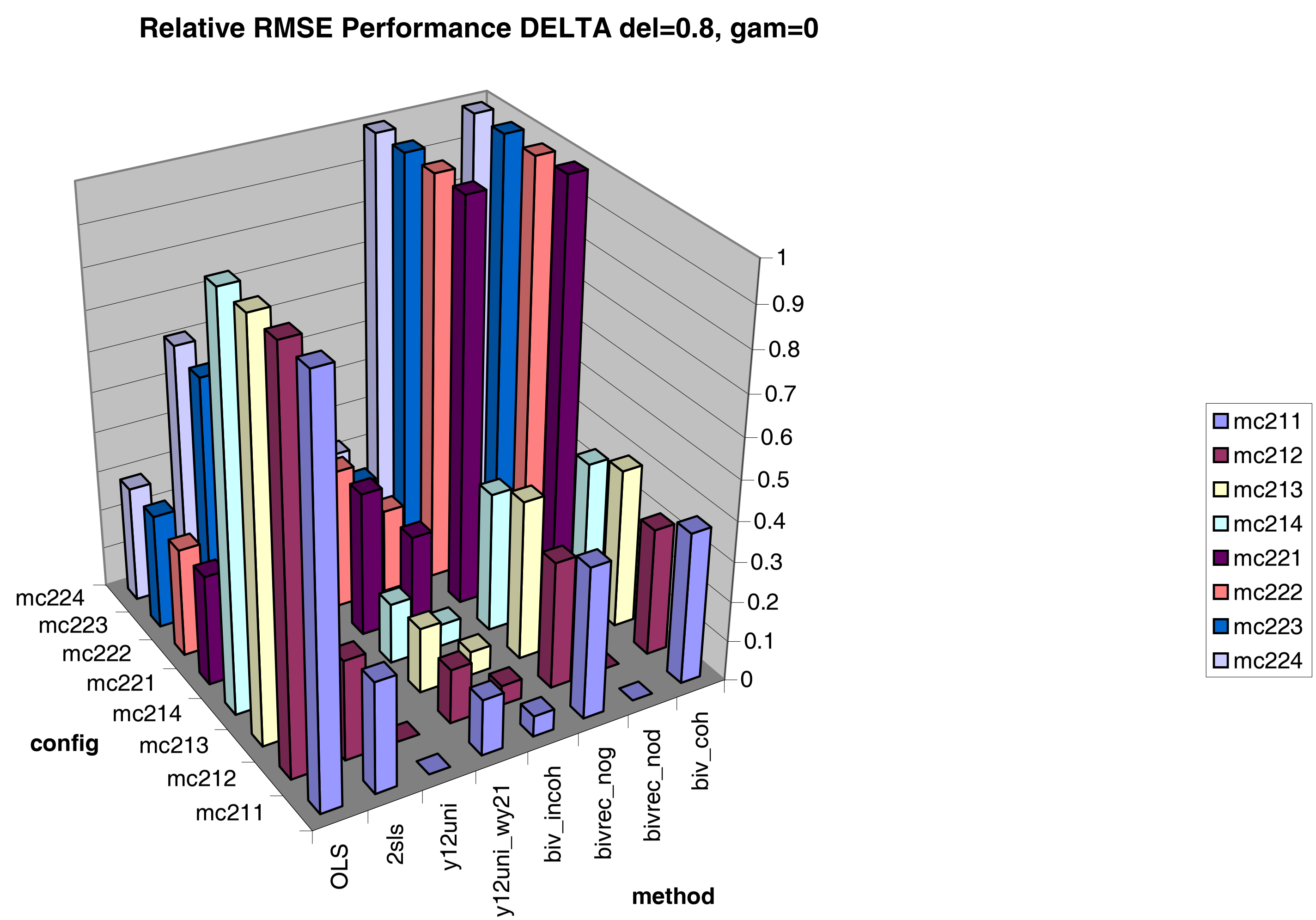
Figure 11



Relative RMSE Performance DELTA del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424

Relative RMSE Performance DELTA del=0.8, gam=1





Relative RMSE Performance DELTA del=0.8, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc224
mc223
mc222
mc221
mc214
mc213
mc212
mc211
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc211
mc212
mc213
mc214
mc221
mc222
mc223
mc224

Relative RMSE Performance DELTA del=0, gam=0

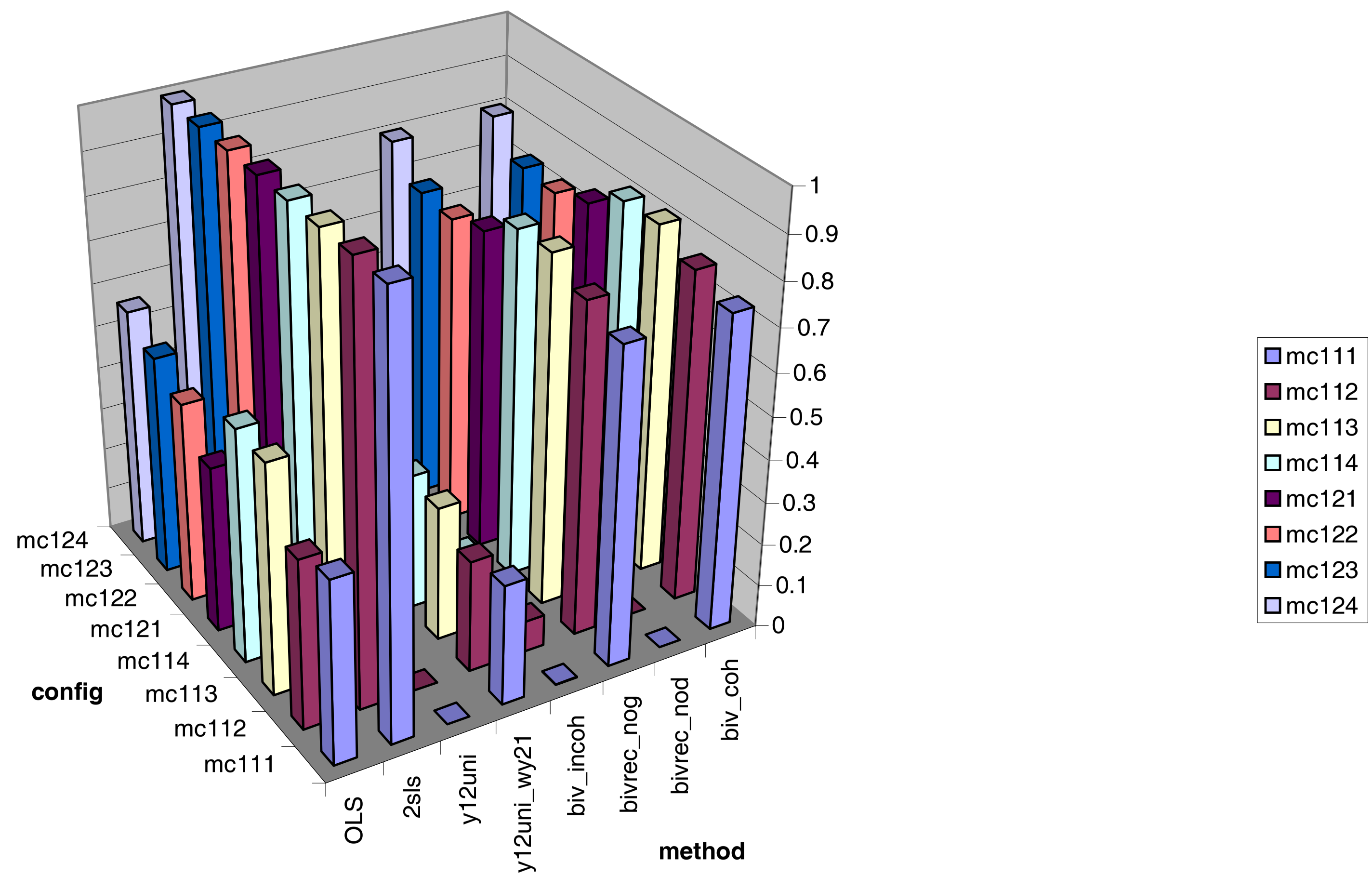
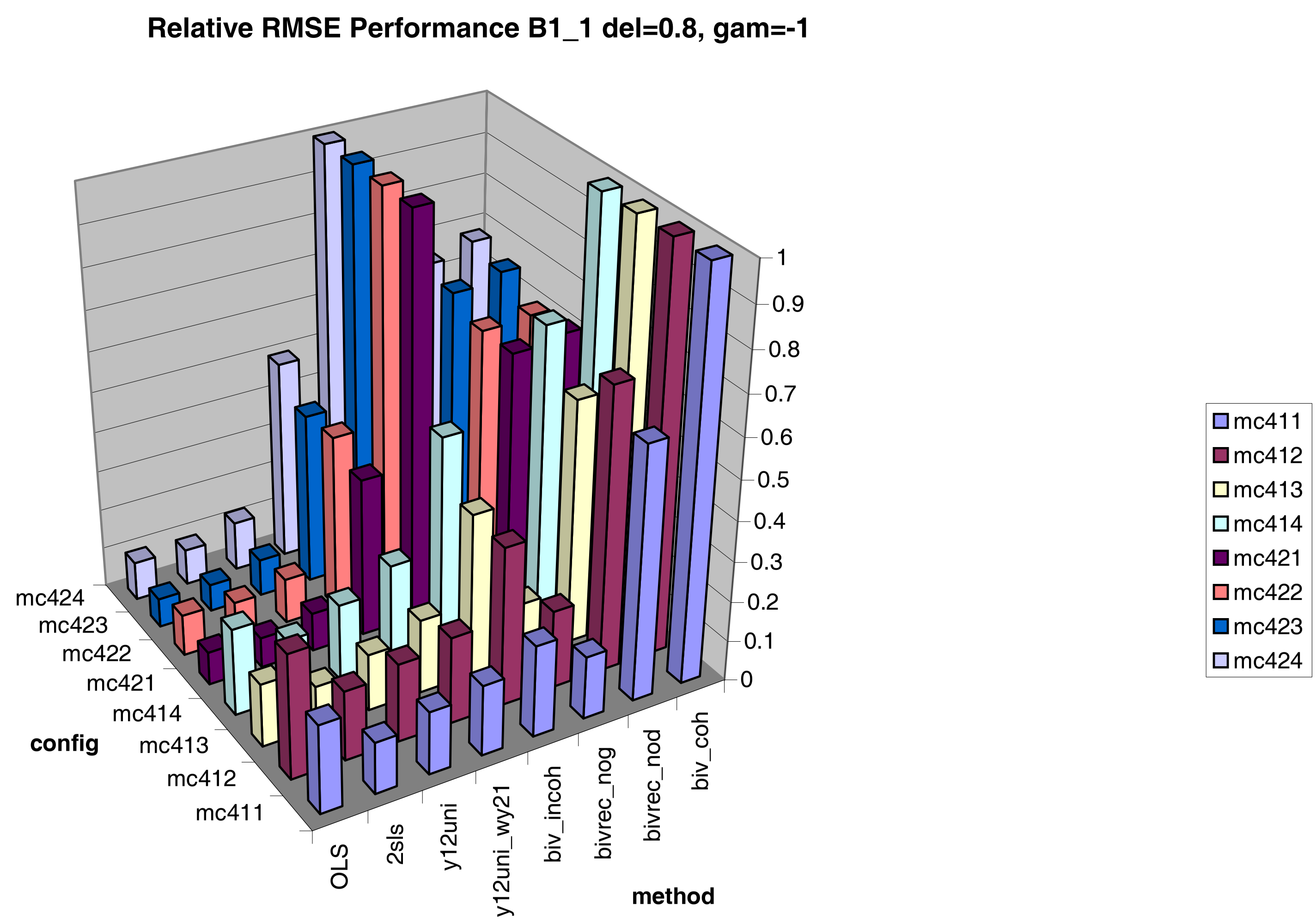
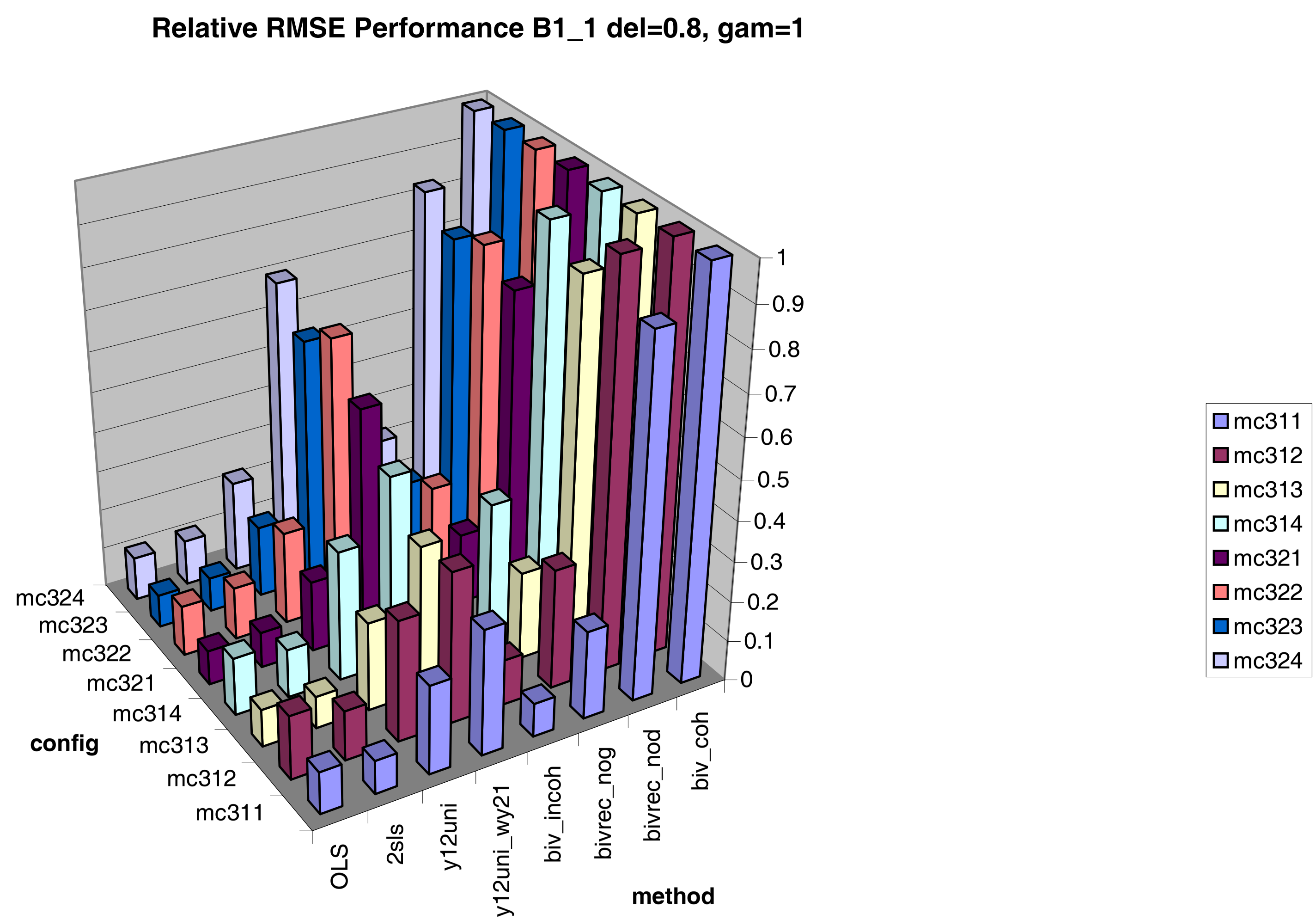
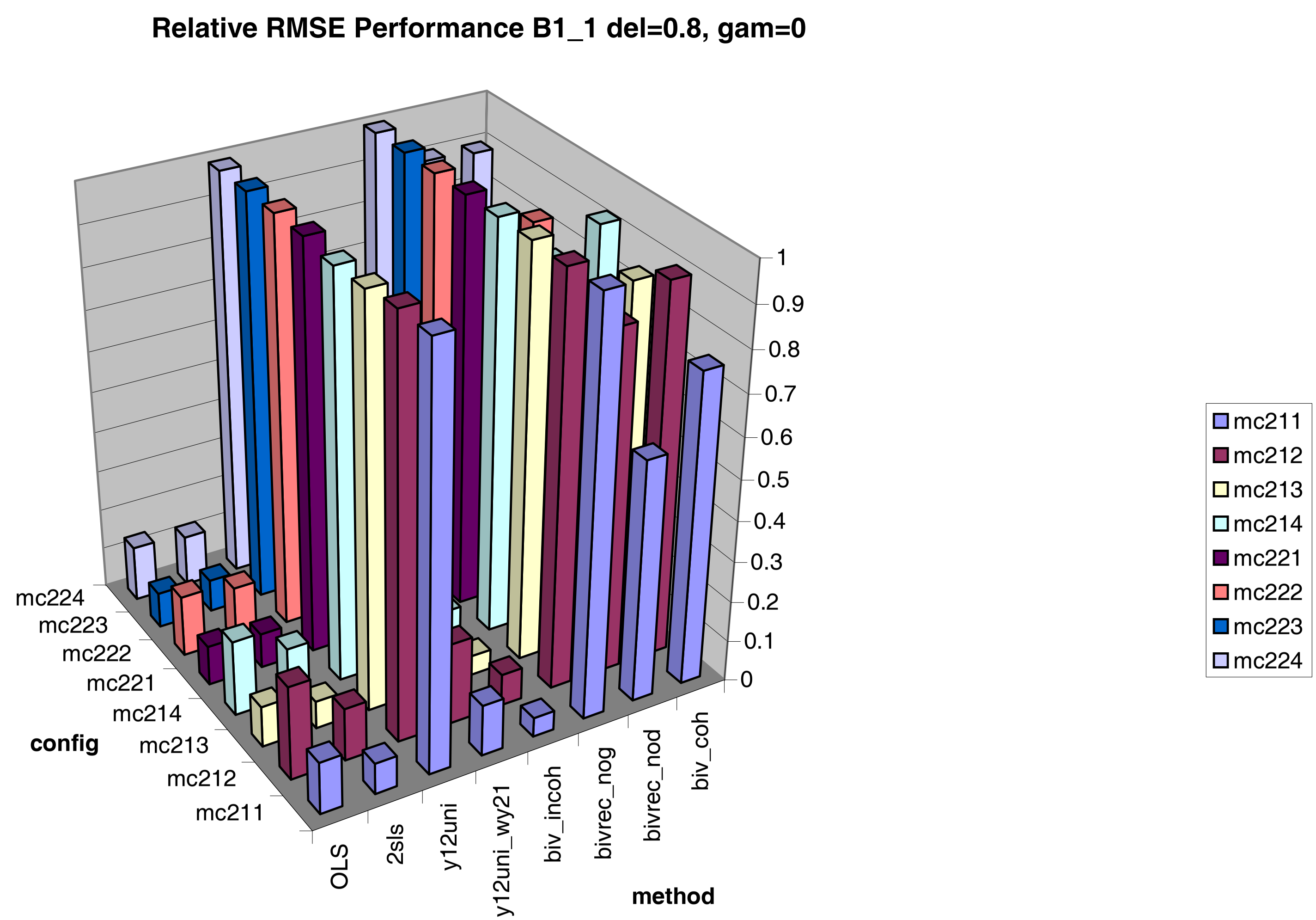


Figure 12



Relative RMSE Performance B1_1 del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424


Relative RMSE Performance B1_1 del=0.8, gam=1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
config
mc311
mc312
mc313
mc314
mc321
mc322
mc323
mc324
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc311
mc312
mc313
mc314
mc321
mc322
mc323
mc324



Relative RMSE Performance B1_1 del=0.8, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc224
mc223
mc222
mc221
mc214
mc213
mc212
mc211
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc211
mc212
mc213
mc214
mc221
mc222
mc223
mc224

Relative RMSE Performance B1_1 del=0, gam=0

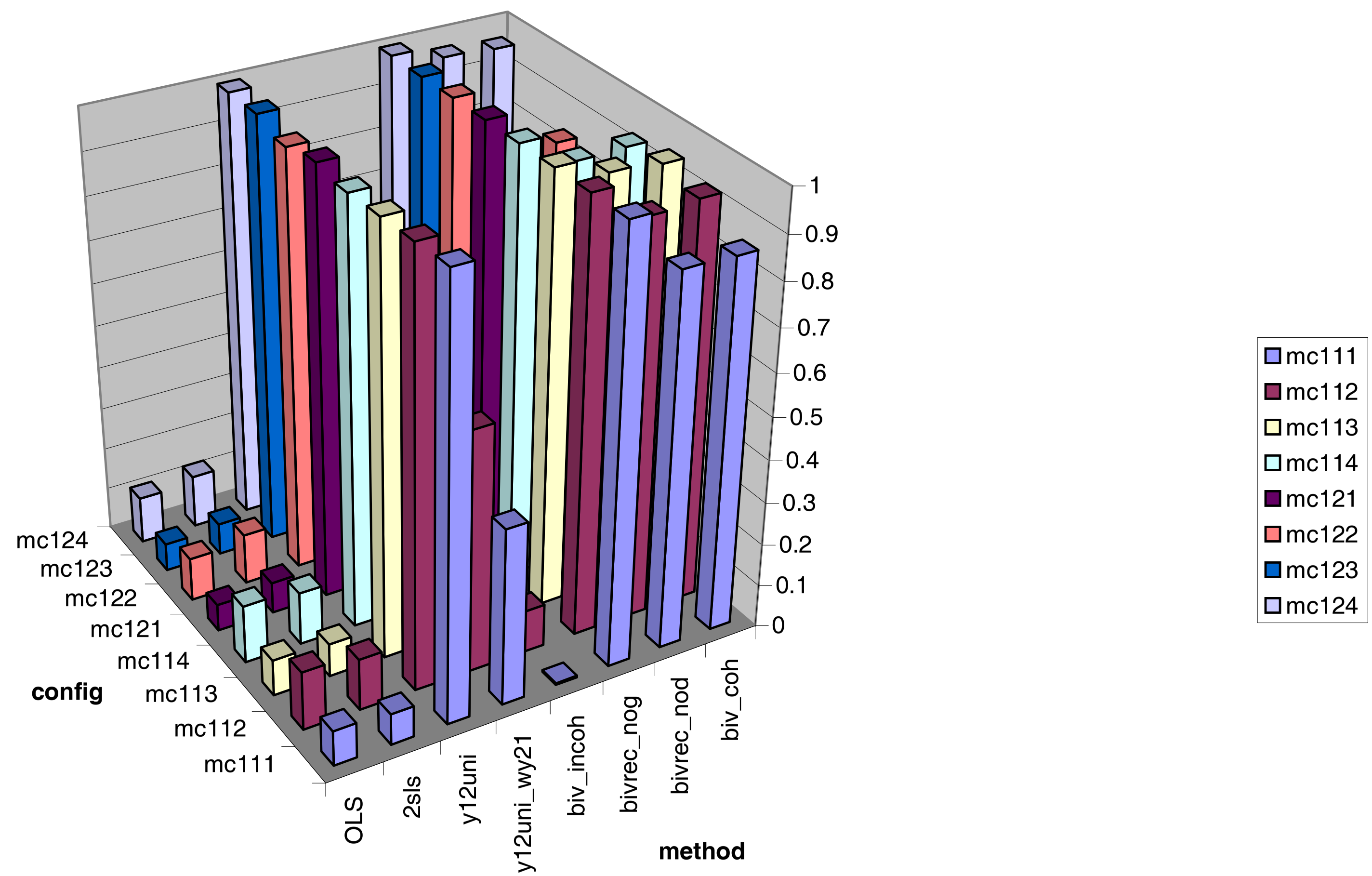
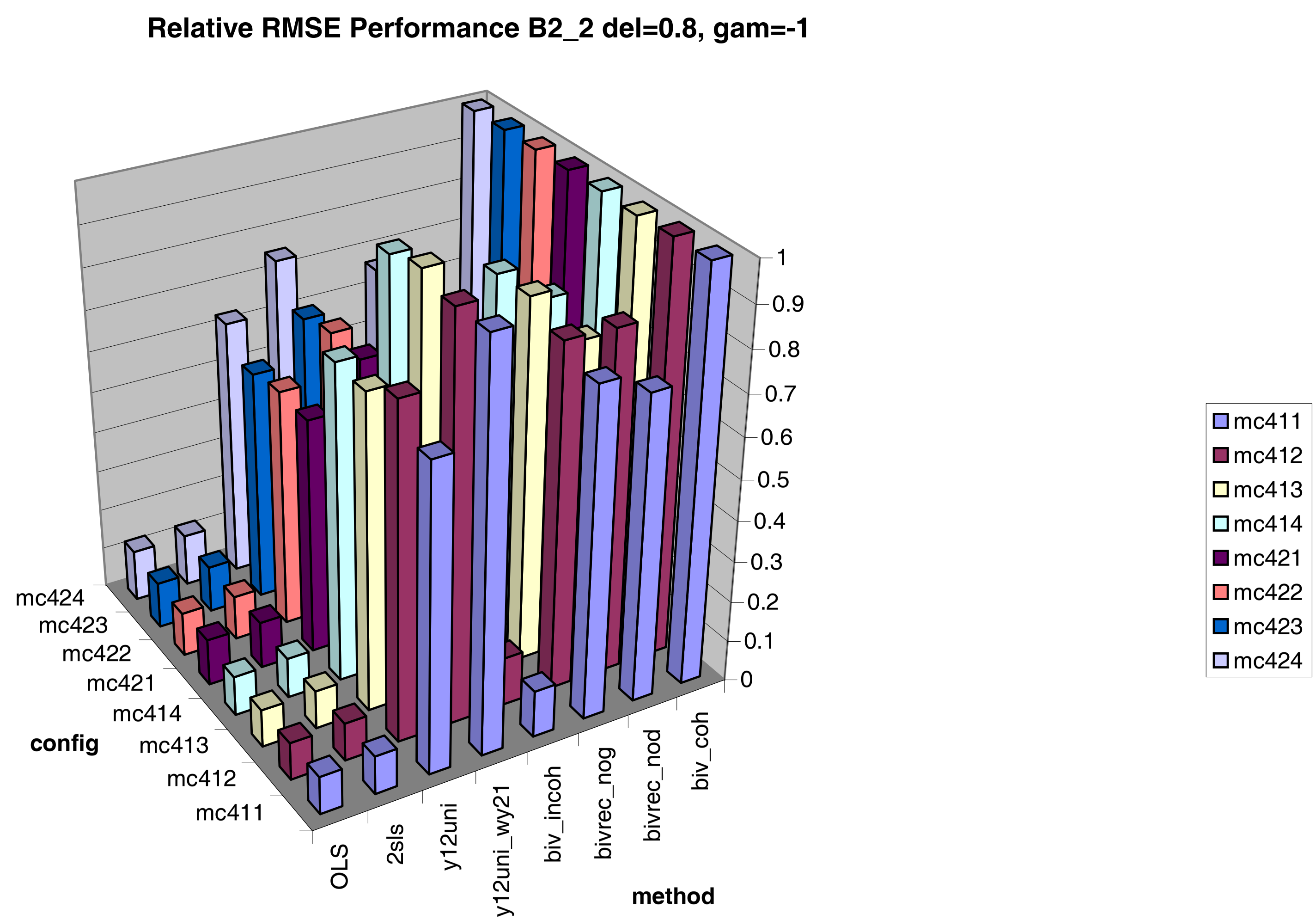
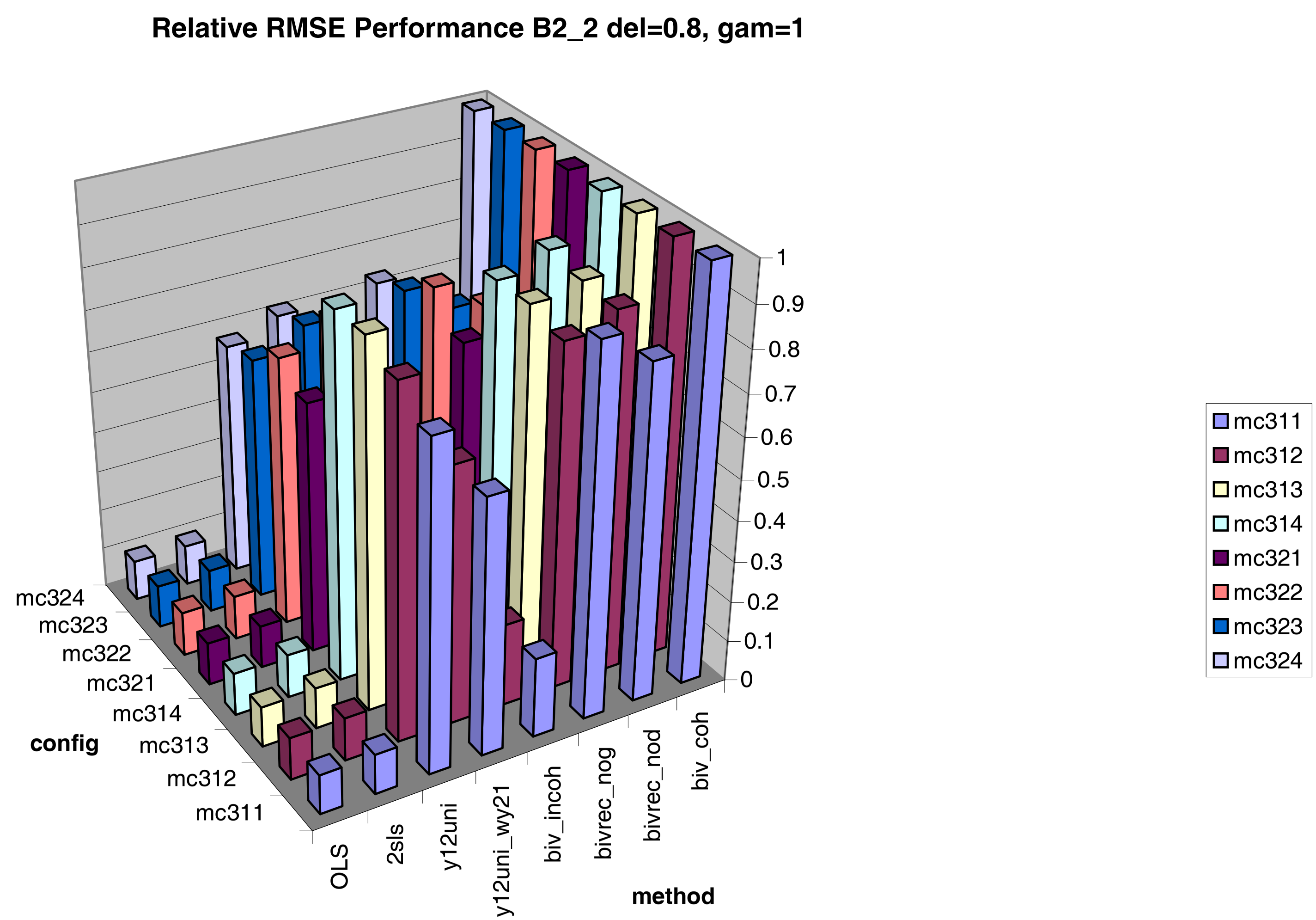
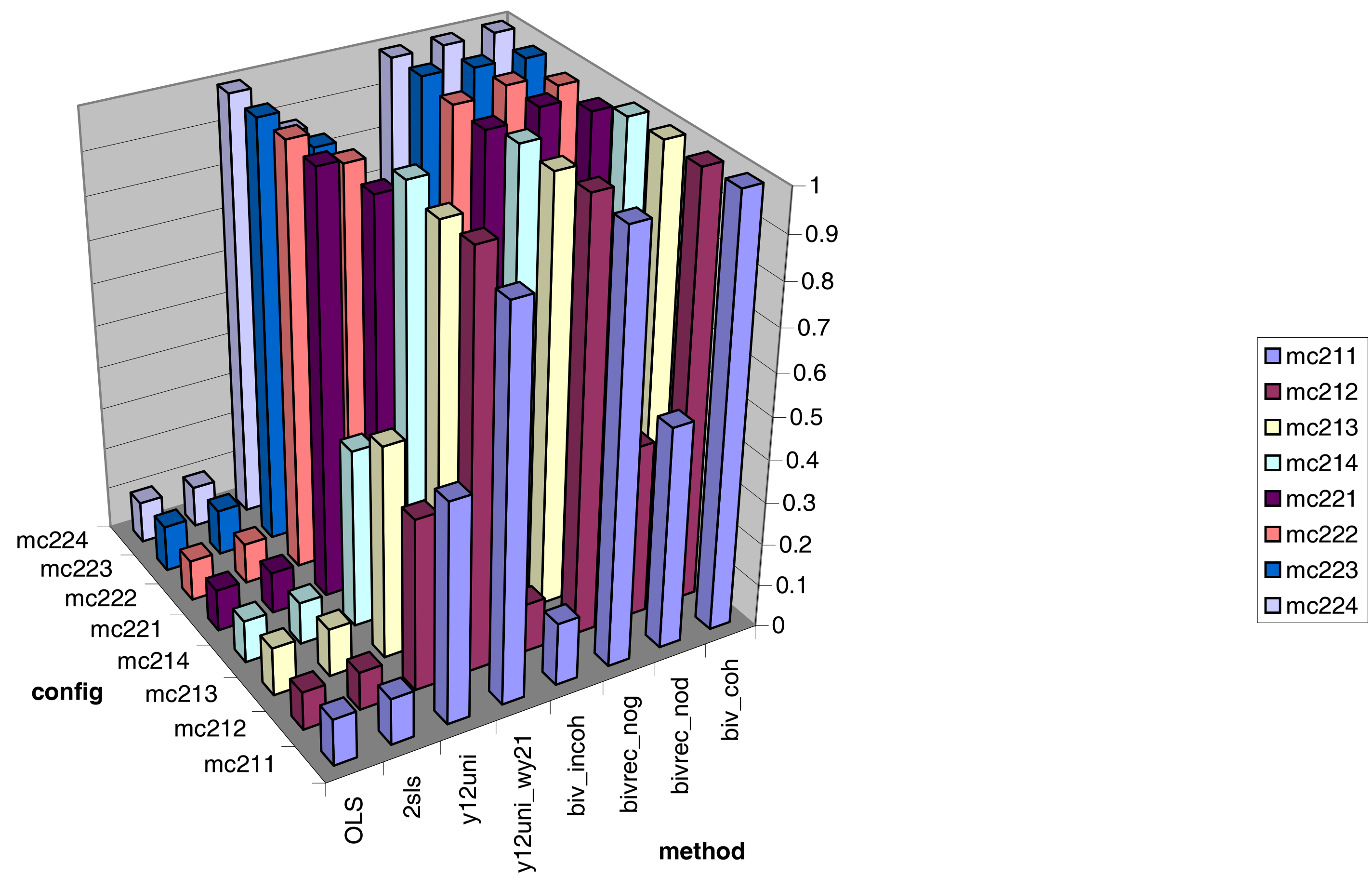


Figure 13


Relative RMSE Performance B2_2 del=0.8, gam=-1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
config
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424


Relative RMSE Performance B2_2 del=0.8, gam=1
mc311
mc312
mc313
mc314
mc321
mc322
mc323
mc324
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Relative RMSE Performance B2_2 del=0.8, gam=0



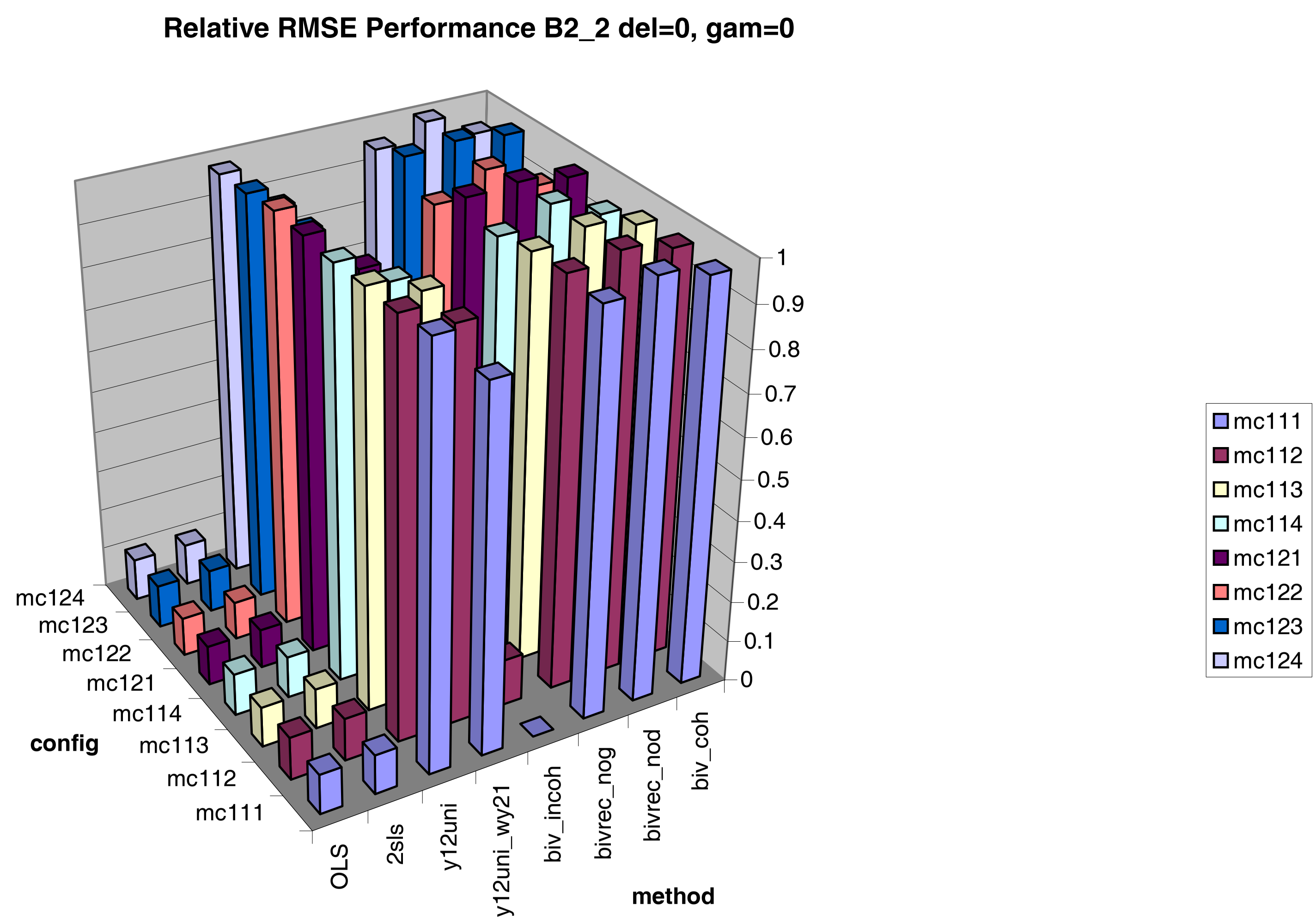
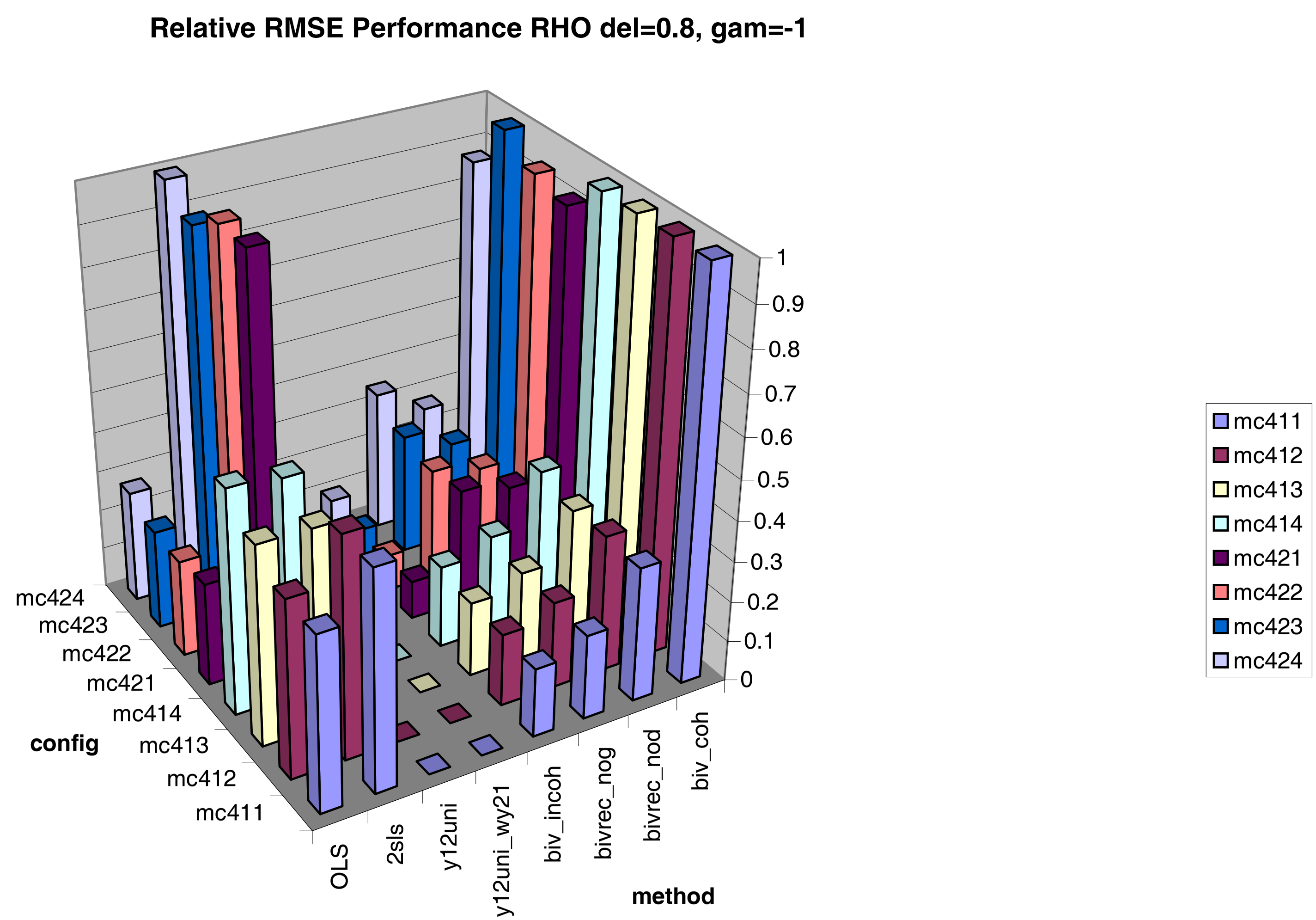
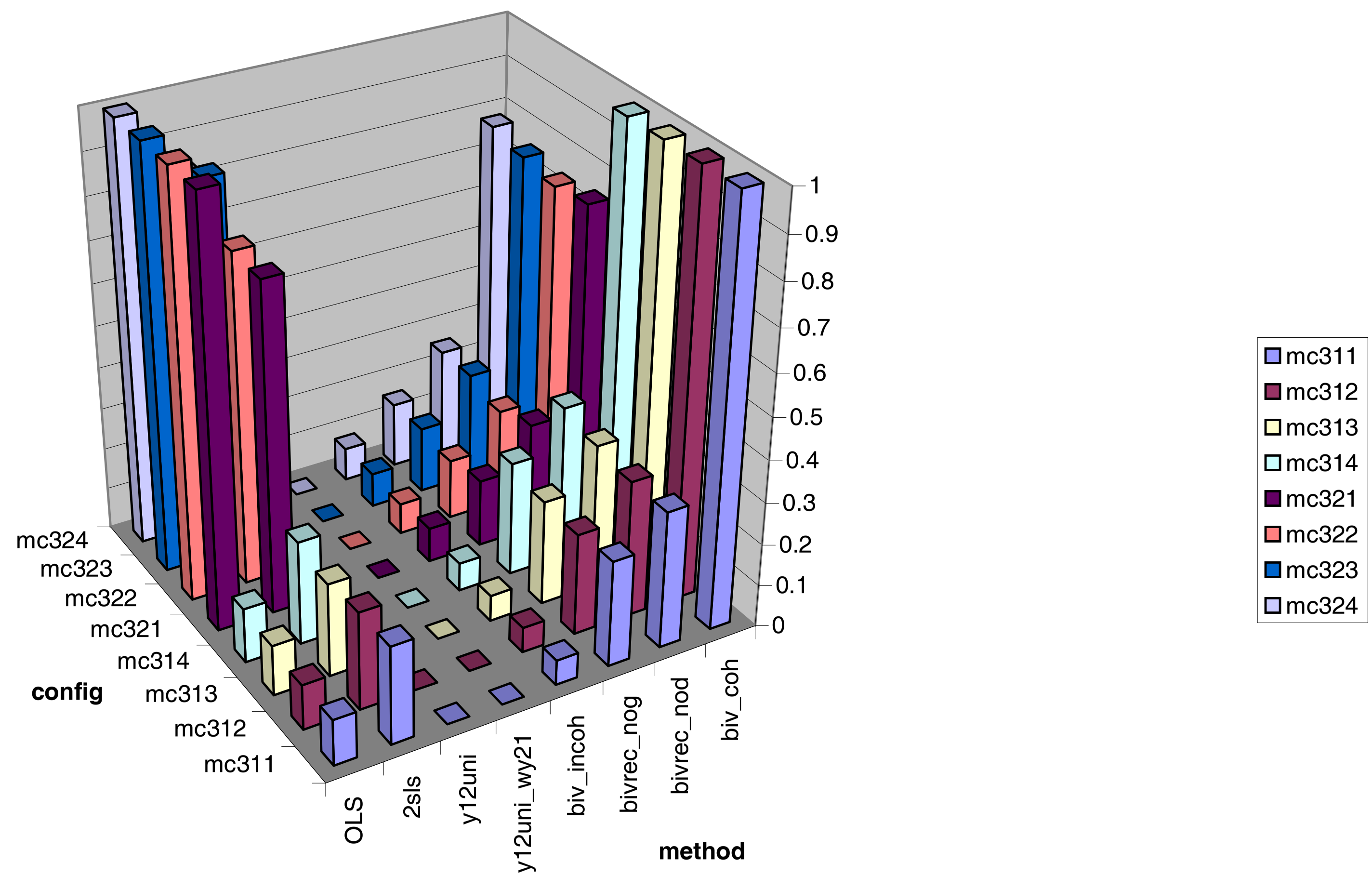

Relative RMSE Performance B2_2 del=0, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
mc124
mc123
mc122
mc121
mc114
mc113
mc112
mc111
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
mc111
mc112
mc113
mc114
mc121
mc122
mc123
mc124

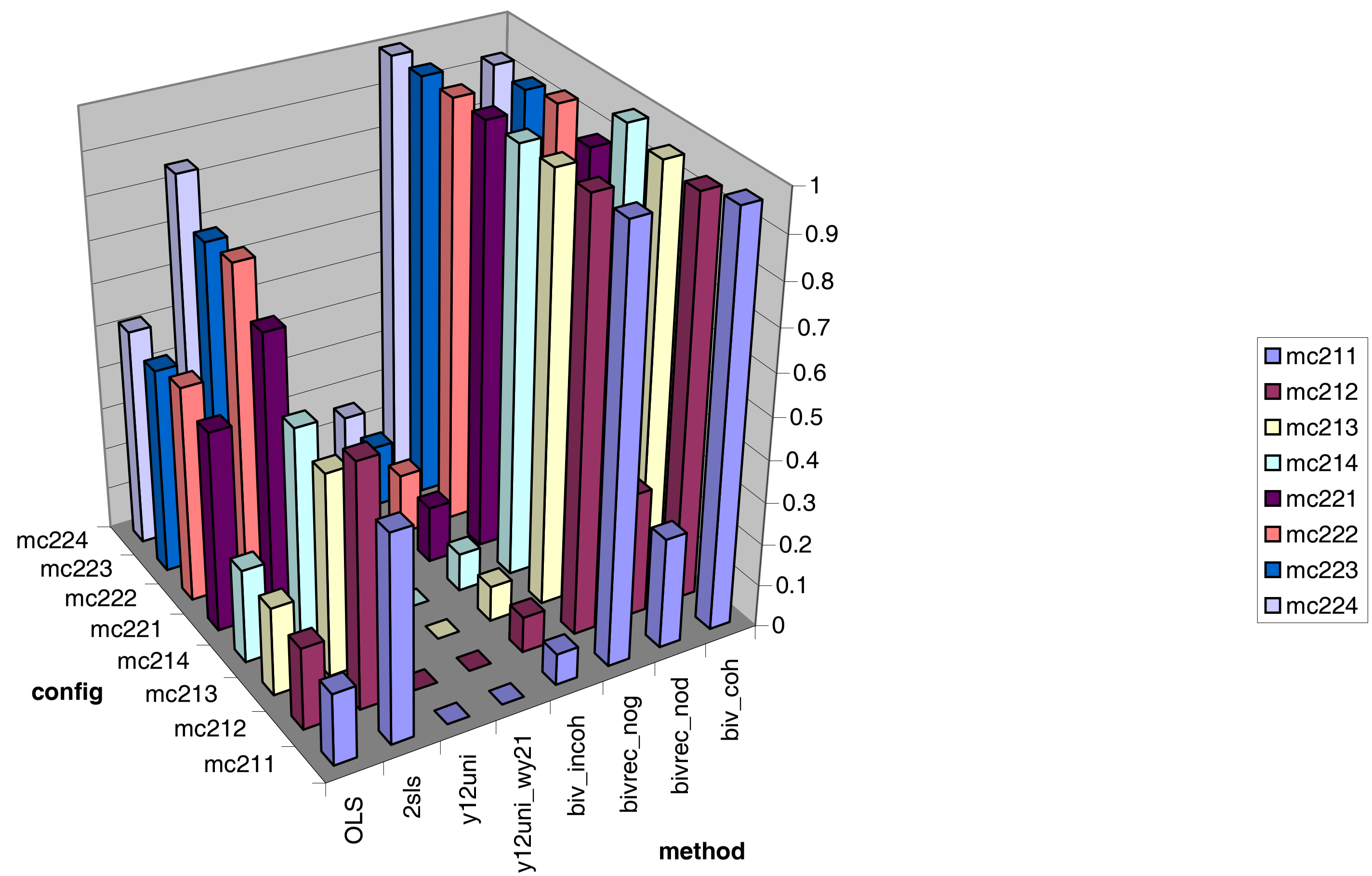
Figure 14


Relative RMSE Performance RHO del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424

Relative RMSE Performance RHO del=0.8, gam=1



Relative RMSE Performance RHO del=0.8, gam=0



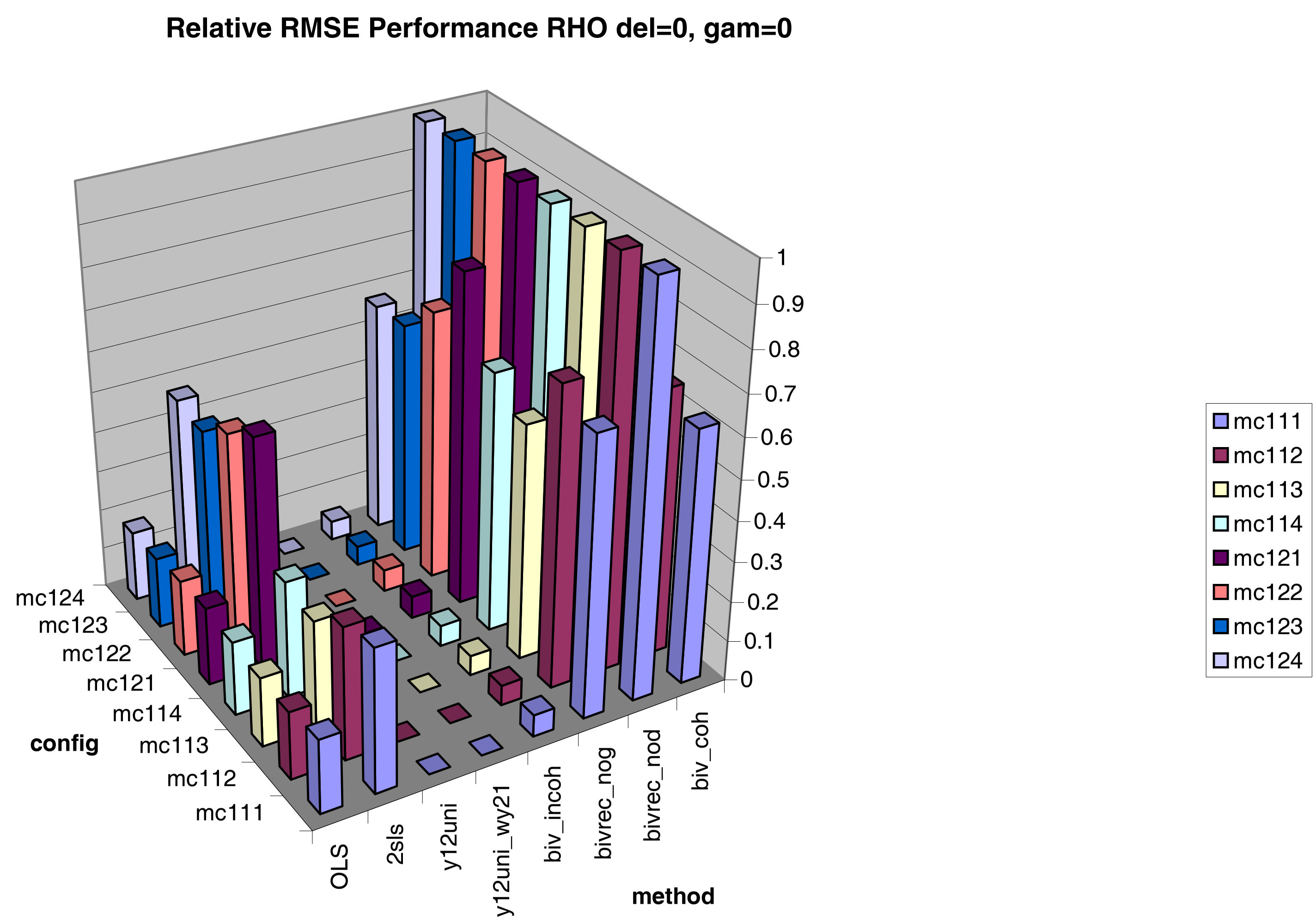
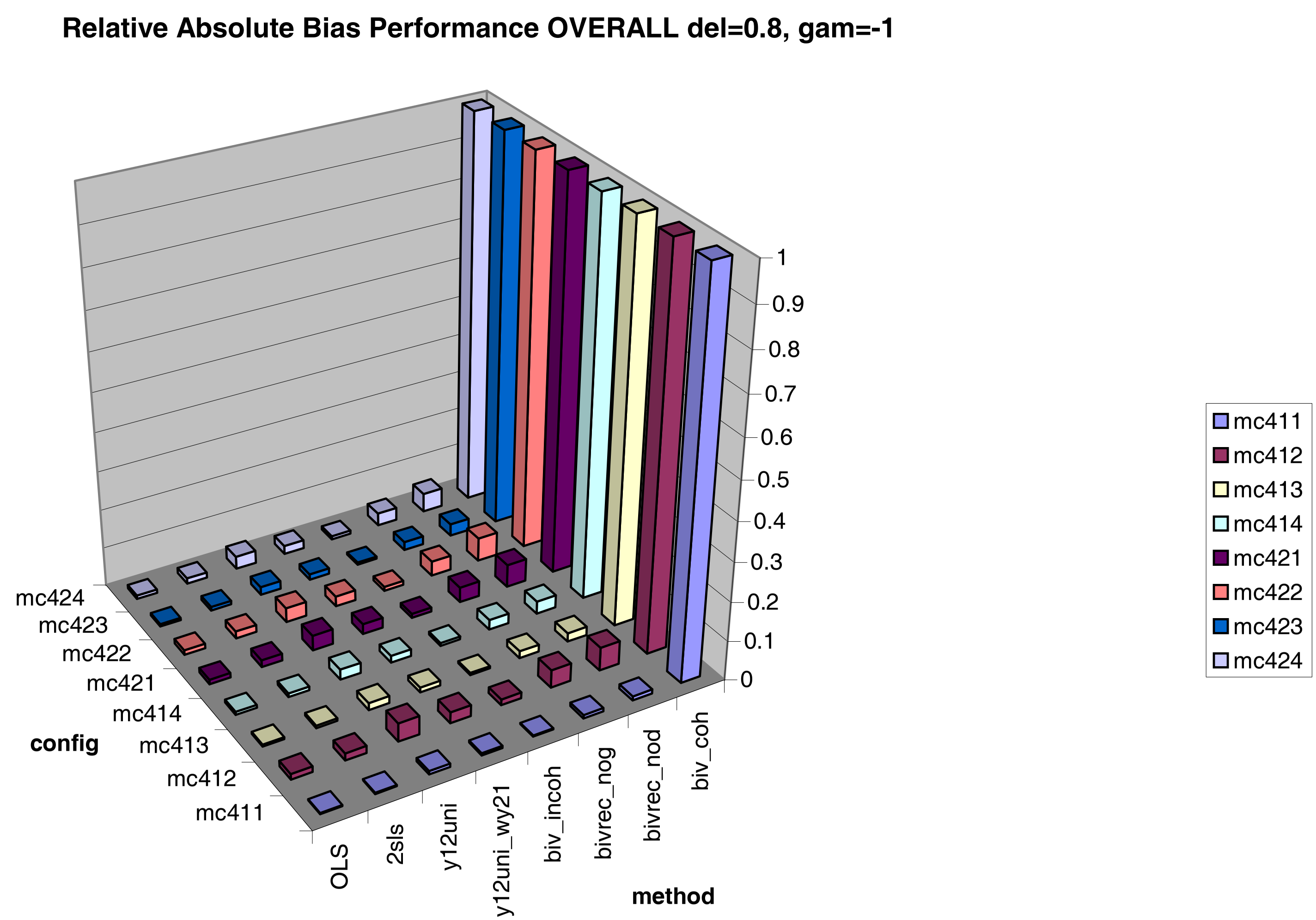
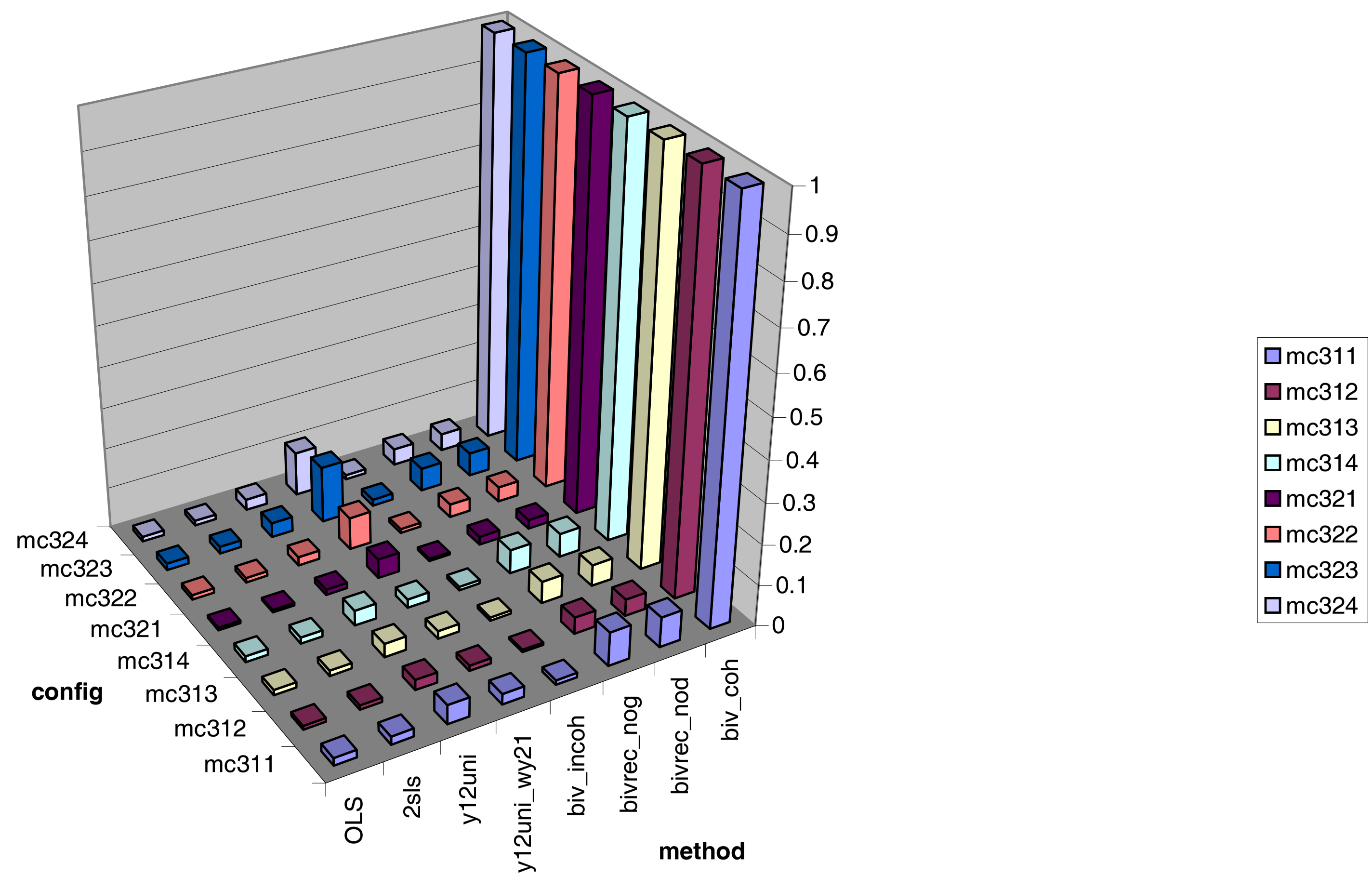

Relative RMSE Performance RHO del=0, gam=0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
mc124
mc123
mc122
mc121
mc114
mc113
mc112
mc111
config
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
method
mc111
mc112
mc113
mc114
mc121
mc122
mc123
mc124

Figure 15

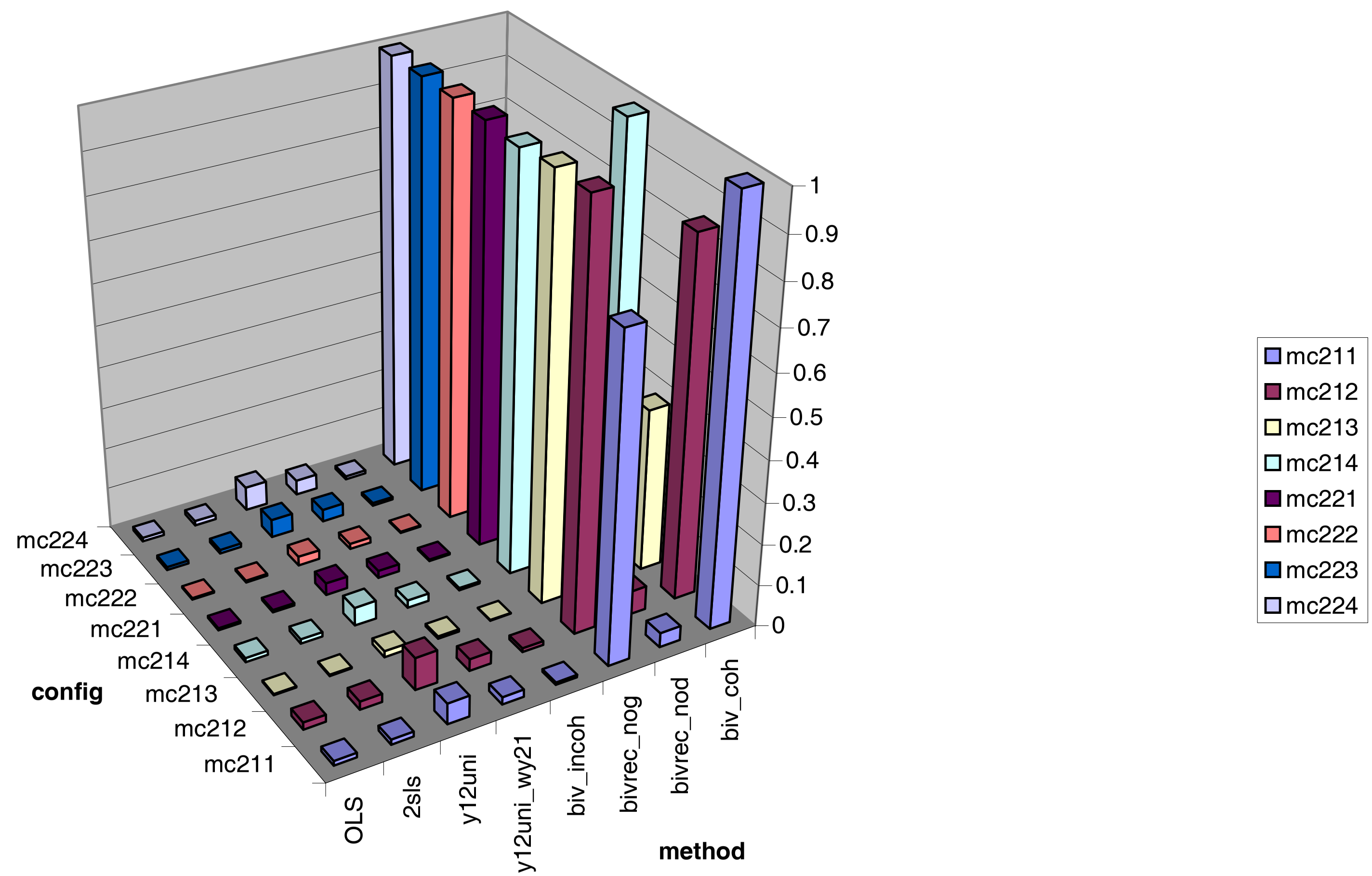


Relative Absolute Bias Performance OVERALL del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424

Relative Absolute Bias Performance OVERALL del=0.8, gam=1



Relative Absolute Bias Performance OVERALL del=0.8, gam=0



Relative Absolute Bias Performance OVERALL del=0, gam=0

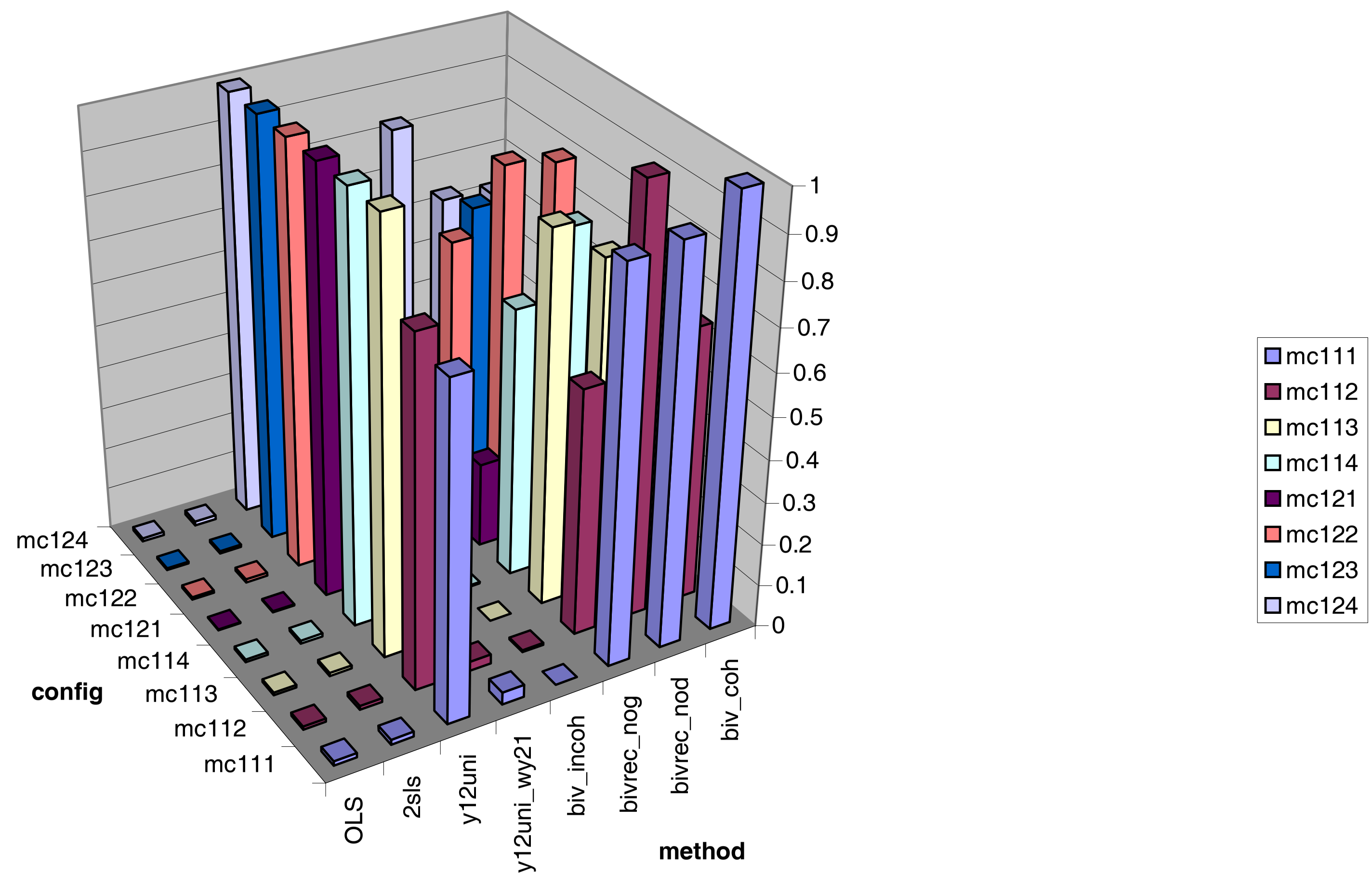
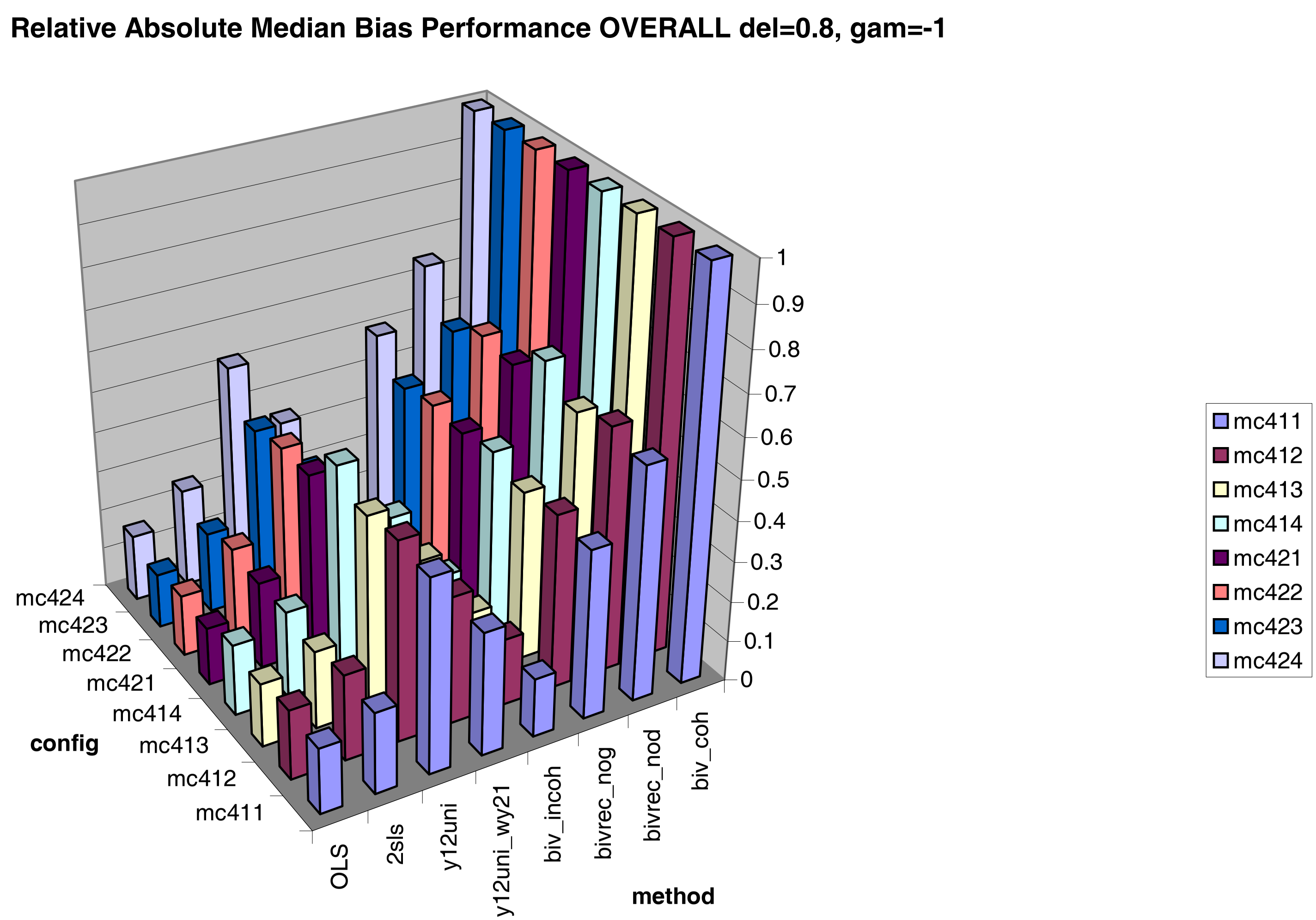
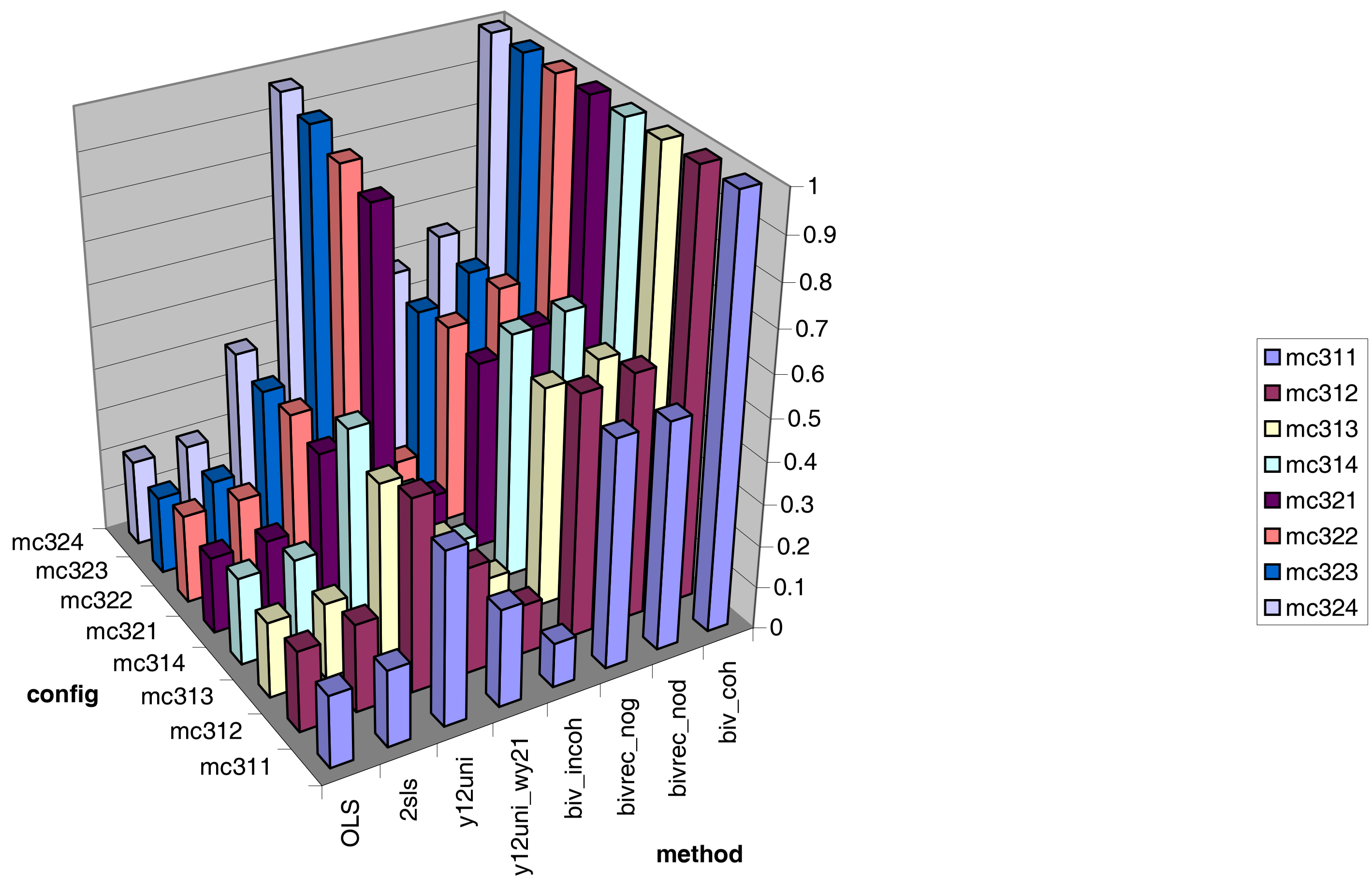


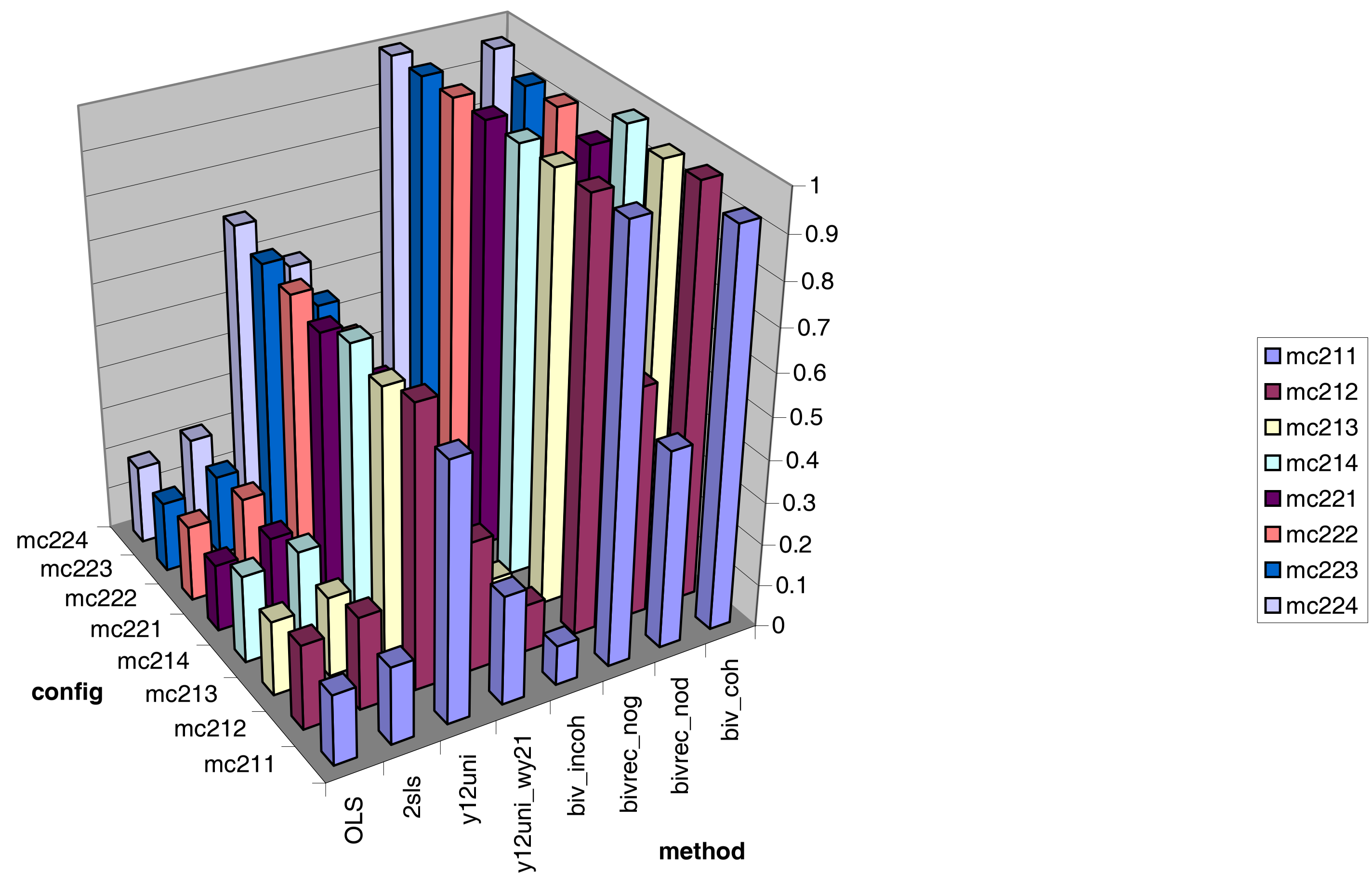
Figure 16


Relative Absolute Median Bias Performance OVERALL del=0.8, gam=-1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
config
mc424
mc423
mc422
mc421
mc414
mc413
mc412
mc411
method
OLS
2sls
y12uni
y12uni_wy21
biv_incoh
bivrec_nog
bivrec_nod
biv_coh
mc411
mc412
mc413
mc414
mc421
mc422
mc423
mc424

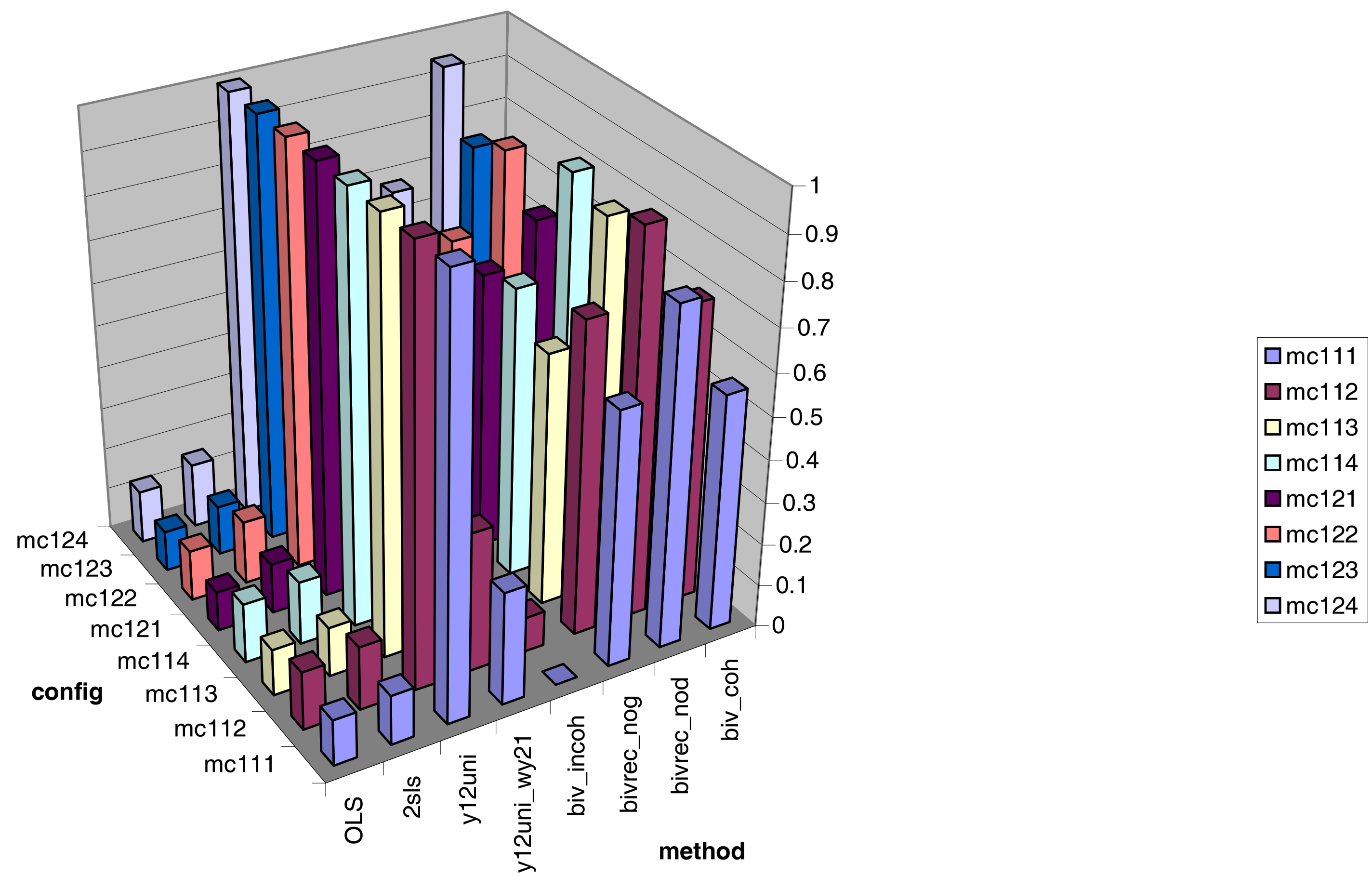
Relative Absolute Median Bias Performance OVERALL del=0.8, gam=1



Relative Absolute Median Bias Performance OVERALL del=0.8, gam=0



Relative Absolute Median Bias Performance OVERALL del=0, gam=0



References

[1] Barghava,A. and J.Sargan (1982): "Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*51, 1635-1660.

[2] Blundell, R. and R. Smith (1994), "Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables," *Journal of Econometrics*, vol.64, pp.355-373.

[3] Dagenais, M. (1997), "A Simultaneous Probit Model," GREQAM working paper No.97A33, University of Aix-en-Marseille.

[4] Falcetti, E. and M. Tudela (2007): "Modelling Currency Crises in Emerging Markets: A Dynamic Probit Model with Unobserved Heterogeneity and Autocorrelated Errors," *Oxford Bulletin of Economics and Statistics*, Volume 68, Number 4, August 2006 , 445-471(27).

[5] Gourieroux C., Laffont J.J, Montfort A. (1980), "Coherency Conditions in Simultaneous Linear Equations Models with Endogenous Switching Regimes," *Econometrica*, 48 (3), pp. 75-95.

[6] Hajivassiliou, V. (2007), "Efficient and Robust Estimation of Panel Data LDV Models with Simultaneity, Dynamics, and Regressor-Random Effects Correlation," September, LSE Department of Economics working paper.

[7] Hajivassiliou V., and Y. Ioannides (2007), "Unemployment and Liquidity Constraints," *Journal of Applied Econometrics*, vol. 22(3), pp.479-510.

[8] Hajivassiliou V., McFadden D.L. (1998), "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica* vol. 66(4), pp.863-896.

[9] Hajivassiliou, V. and F. Savignac (2007): "Financing Constraints and a Firm's Decision and Ability to Innovate: Establishing Direct and Reverse Effects," LSE FMG working paper No.594.

[10] Heckman, J. (1978), "Dummy Endogenous Variables in Simultaneous Equation Systems," *Econometrica* vol.66, pp.931-960.

[11] Heckman, J (1981): "Statistical Models for Discrete Panel data," in Manski and McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, Cambridge, Massachusetts.

[12] Heckman, J. (1981b): "Dynamic Discrete Models," in Manski and McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, Cambridge, Massachusetts.

[13] Lewbel A. (2007), "Coherency and Completeness of Structural Models Containing a Dummy Endogenous Variable," *International Economic Review* 48(4), November, pp.1379-1392.

[14] Maddala, G. and L.-F. Lee (1976), "Recursive Models with Qualitative Endogenous Variables," *Annals of Economic and Social Measurement*, vol.5, pp.525-545."

[15] McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica* 57, 995-1026.

[16] Tamer, E. (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies* 70, 147-165.

[17] Uhlig H. (2005), "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, vol.52(2), pp. 381-419.