



Econométrie des coûts de production agricole : comparaison de deux méthodes d'estimation sur la base du RICA





L'estimation économétrique des coûts de production agricole : objectifs

- *Expliquer le développement inégal des différentes spéculations agricoles*
- *Etudier l'adaptation des agriculteurs à la réforme de la Politique agricole communautaire (PAC)*
- *en*
 - *analysant les coûts de production des principaux produits agricoles européens*
 - *estimant la répartition des charges entre les différentes productions*
 - *utilisant les données comptables des exploitations agricoles*



*Le modèle coûts de production INRA-INSEE-SCEES :
les principales étapes de son élaboration*

- *Estimation micro-économique d'un tableau d'input-output (Divay & Meunier 1980)*
- *Les coûts de production des grands produits agricoles (Aufrant 1983)*
- *L'agriculture dans la CEE : coûts de production (Butault & alii 1990)*
- *Le modèle sur les coûts de production agricole (Pollet & alii 1996)*
- *Les coefficients de marge brute standard 1996 (Demotes-Mainard & alii, 2001)*



Le modèle Coûts de production

Application - Variables explicatives : les produits

- *les 25 produits :* $X_j \quad j = 1, \dots, 25$
en colonne

blé tendre , blé dur, orge, maïs, autres céréales, légumes secs, pommes de terre, betteraves sucrières , colza, tournesol, tomates, fruits à noyaux, fleurs, fruits à pépin, vin de table, vin de qualité, autres végétaux, bovins, ovins, porcins, volailles, lait de vache et produits laitiers bovins, lait de brebis et produits laitiers ovins, oeufs, autres produits animaux.

- **Le concept de produit brut est pris hors intra-consommation**
Produits bruts = Ventes + Variation_de_stocks + Autoconsommation



Le modèle Coûts de production

Application - Variables expliquées : les charges

■ *Les 15 postes de charges*

en ligne :

$$Y_k \quad k = 1, \dots, 15$$

- *charges proportionnelles : aliments du bétail, frais spécifiques élevage, semences et plants, engrais et amendements, produits phyto-sanitaires, produits pétroliers,*
- *charges fixes : entretien bâtiments et matériel, coûts fonciers, taxes (hors foncier), intérêts (hors foncier), amortissements foncier et bâtiments, amortissements matériel, assurance, autres biens et services*
- *subventions (charge négative)*

■ **Le concept de coût retenu est celui de charges réelles**

**Charges réelles = Consommations_intermédiaires +
Frais_d'exploitation + Amortissements**



Le modèle Coûts de production

Application - Variables expliquées : le revenu

Le solde du revenu en ligne : $Y_{16} = R$ $k = 16$

Le concept de revenu hors travail de l'exploitation

revenu net hors travail = produit brut + subventions -
somme des charges fixes et proportionnelles

$$R = \sum_{j=1}^p (P_j - C_j) \times Q_j$$

avec

- P_j prix unitaire du bien j
- C_j coût unitaire du bien j
- Q_j quantités produites du bien j



Le modèle Coûts de production

Méthodologie - Hypothèses et spécification

Méthodologie : reconstitution d'une matrice d'input-output à partir de données micro-économiques (Divay et Meunier, 1980)

Hypothèses :

- *H1 - même technique de production pour toutes les exploitations à un facteur aléatoire près $Y_k = f(X_j) + U_k$*
- *H2 - la valeur d'input utilisée est proportionnelle à la valeur de la production $f(X_j) = A'_j * X_{Ji}$*

Spécification du modèle :

Avec : Y_k - Coût de l'input k

a_{jk} - Coefficient technique

X_j - Produit brut du bien j

U_k - Résidu aléatoire iid

$$Y_k = \sum_{j=1}^p a_{jk} * X_j + U_k$$



Le modèle Coûts de production Méthodologie - Hypothèses et spécification

Hypothèses :

- H3 - les U_j sont d'espérance nulle, identiquement et indépendamment distribués
- H4 - le revenu en un bien j est proportionnel à sa production $R_j = b_j * X_j$

Spécification du modèle :

$$R = \sum_j b_j * X_j + \varepsilon ; \varepsilon \text{ iid}$$

le revenu est fonction linéaire de la production

Contrainte de cohérence :

$$\sum_k a_{jk} + b_j = 1 \quad \forall j$$

la production en un bien est la somme des coûts et du revenu



***Le modèle Coûts de production :
les problèmes d'estimation rencontrés***

***Modèle initial : régression des MCO en coupe annuelle
avec procédure de redressement des estimateurs***

- ***variance forte des estimateurs***
- ***coefficient de régression négatifs***
- ***terme constant significativement différent de zéro***
- ***coefficients sous-estimés ou instables dans le temps***

- 1) Identifier l'origine de ces problèmes d'estimation***
- 2) Proposer des méthodologies susceptibles d'améliorer les estimations***
- 3) Evaluer l'intérêt pratique des méthodologies proposées***



***Problèmes d'estimation :
problèmes de précision des estimateurs***

■ Origine :

- hétéroscédasticité pour chacune des k charges (équations)

$$E(U_{ki}, U_{ki'}) \neq 0$$

- dispersion des résidus variant en fonction du niveau de production

$$\text{var}(u_{kji}) \neq \sigma^2 \quad i = 1, \dots, N$$

$$\text{cov}(U_k) \neq \sigma^2 I_N$$

- les estimations MCO menées pour chacune des K charges ne sont pas de variance minimale

■ solutions:

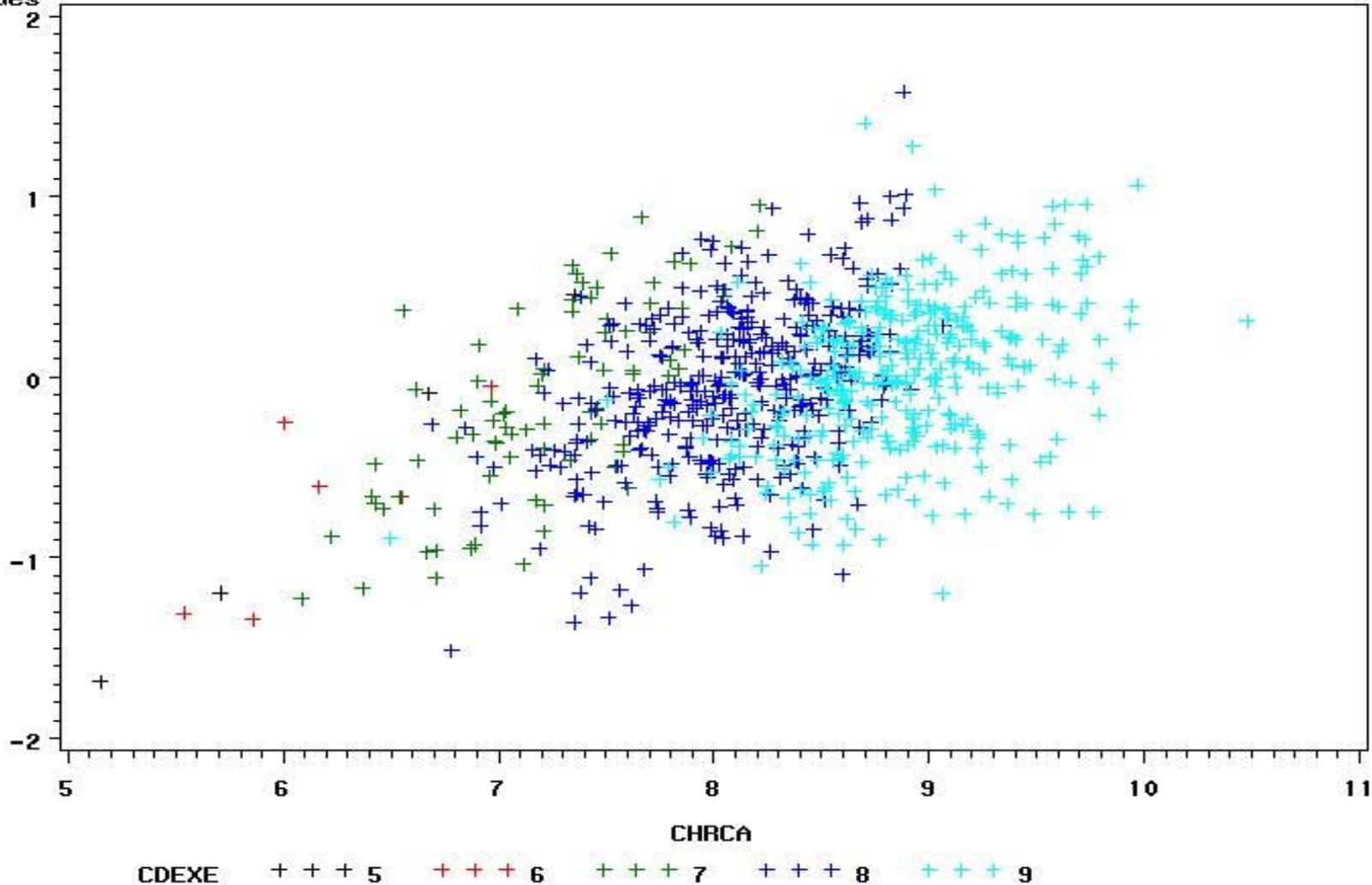
- les procédures de moindres carrés généralisées sont préférables

Econométrie des coûts de production agricole

Problèmes d'estimation : dispersion des résidus en fonction de la dimension de l'exploitation



Residual Values





***Problèmes d'estimation :
problèmes de précision des estimateurs***

■ Origine :

- corrélation entre résidus de certaines équations de charges

$$\text{cov}(U_k, U_{k'}) \neq 0$$

■ solutions:

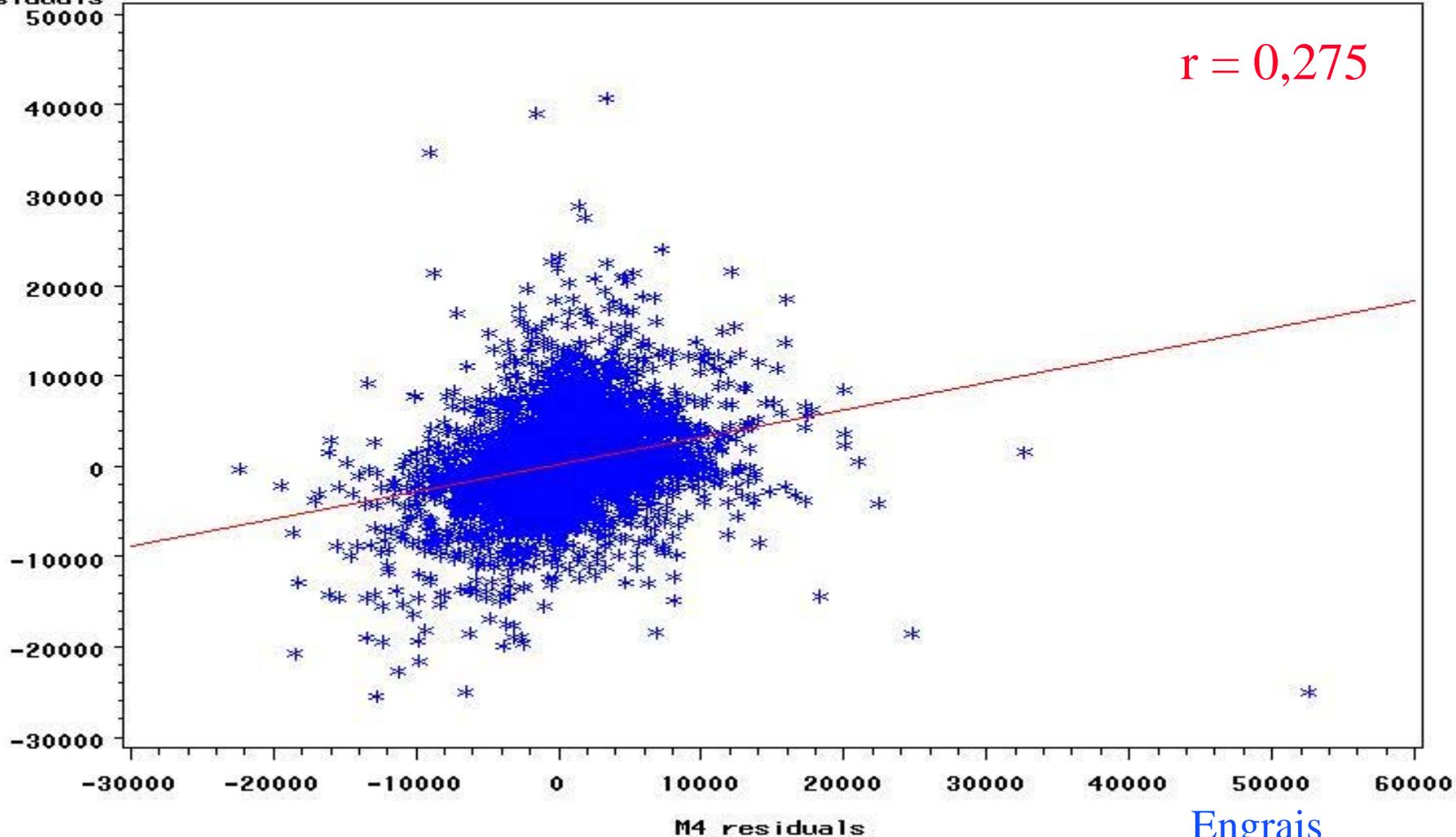
- les procédures d'estimation de modèles à équations simultanées par des méthodes de moindres carrés généralisés (estimation SURE) permettent de prendre en compte les corrélations entre résidus des équations de charge

Econométrie des coûts de production agricole
Problèmes d'estimation : corrélations entre résidus
de plusieurs équations



Phytosanitaires

M5 residuals



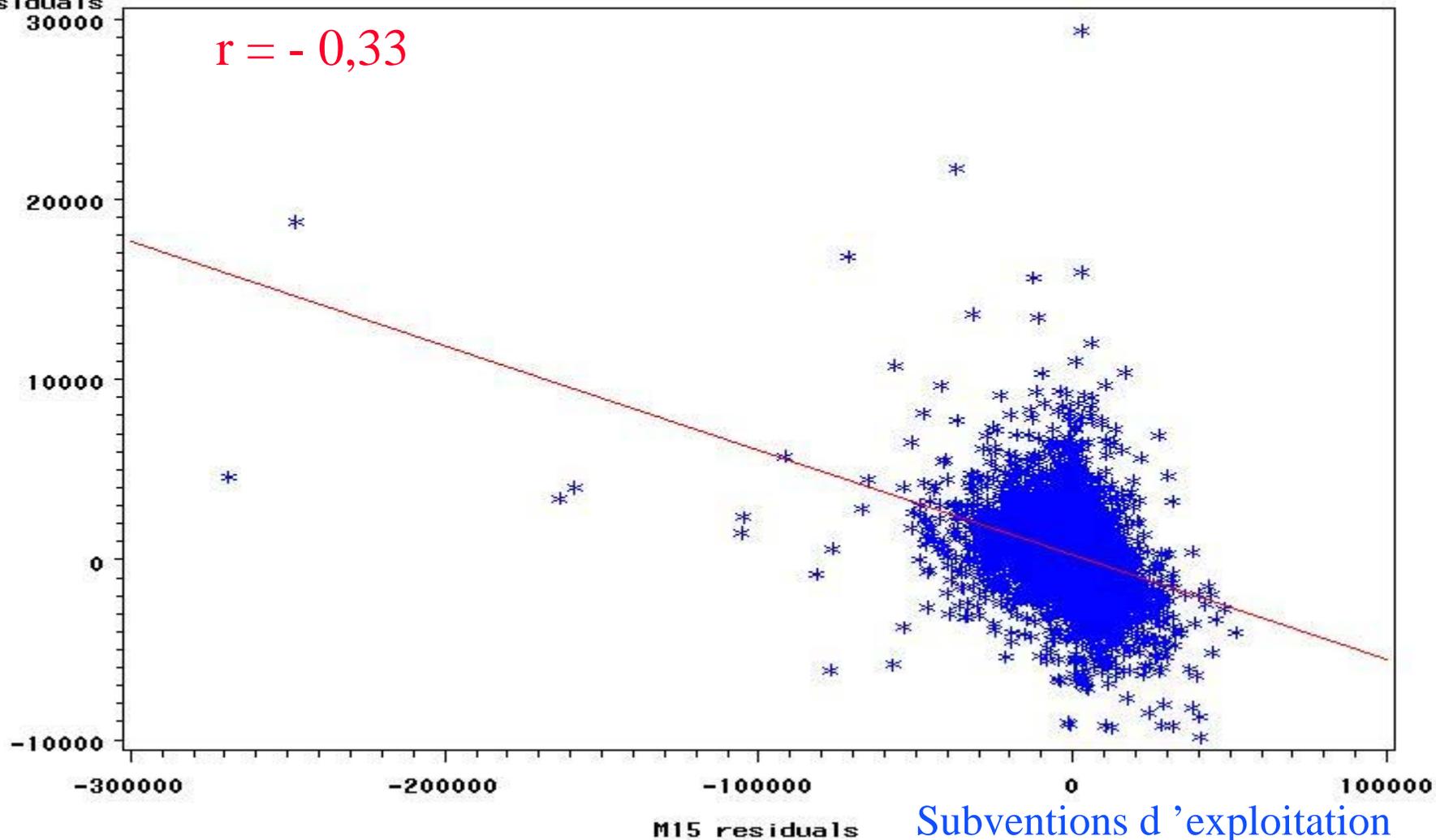
Engrais

Econométrie des coûts de production agricole
Problèmes d'estimation : corrélations entre résidus
de plusieurs équations



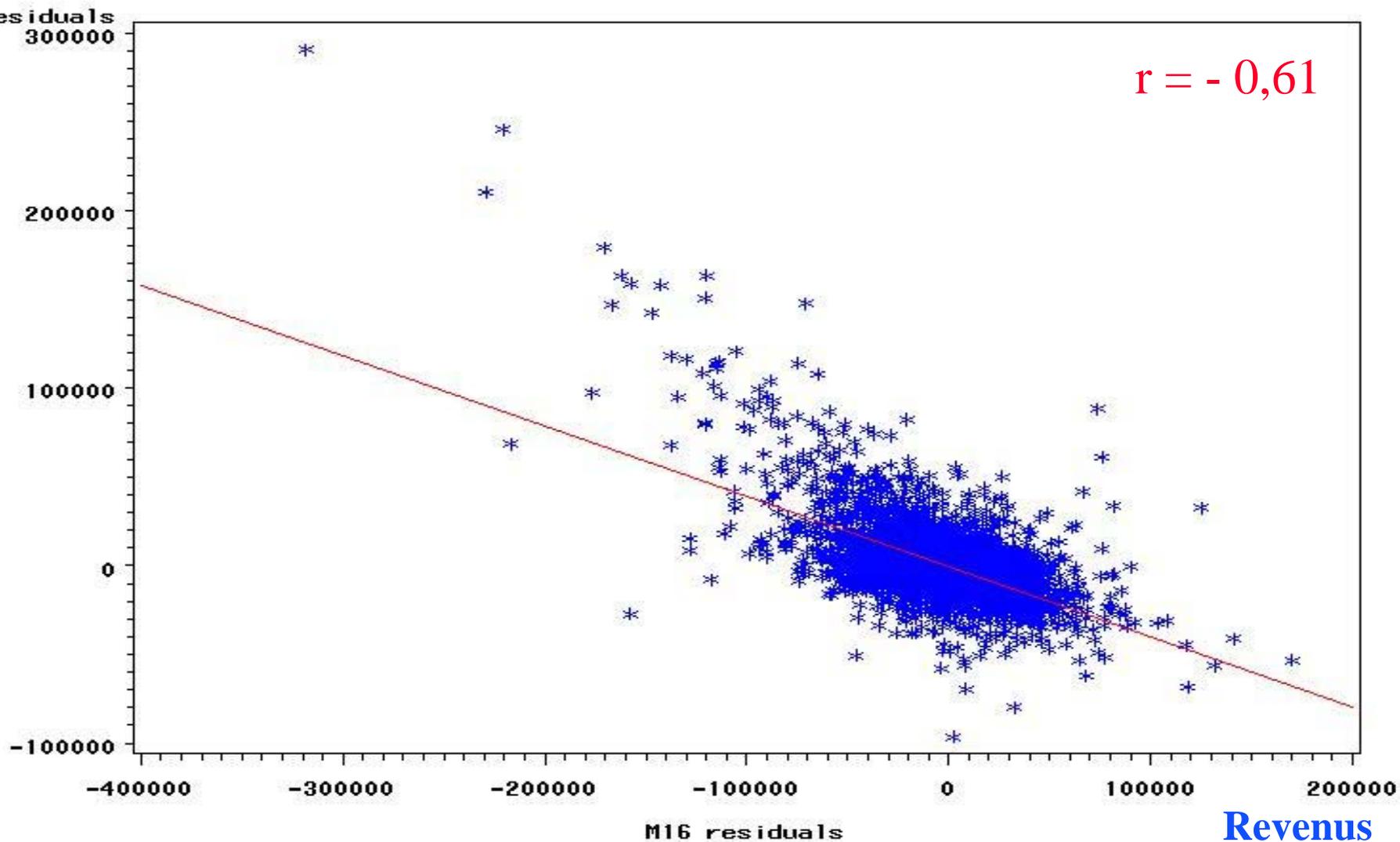
Assurances

M13 residuals





Autres biens et services





*Problèmes d'estimation :
problèmes de précision des estimateurs*

■ Origine :

- biais de simultanéité des estimateurs MCO

$$\text{cov}(U_k, X_j) \neq 0$$

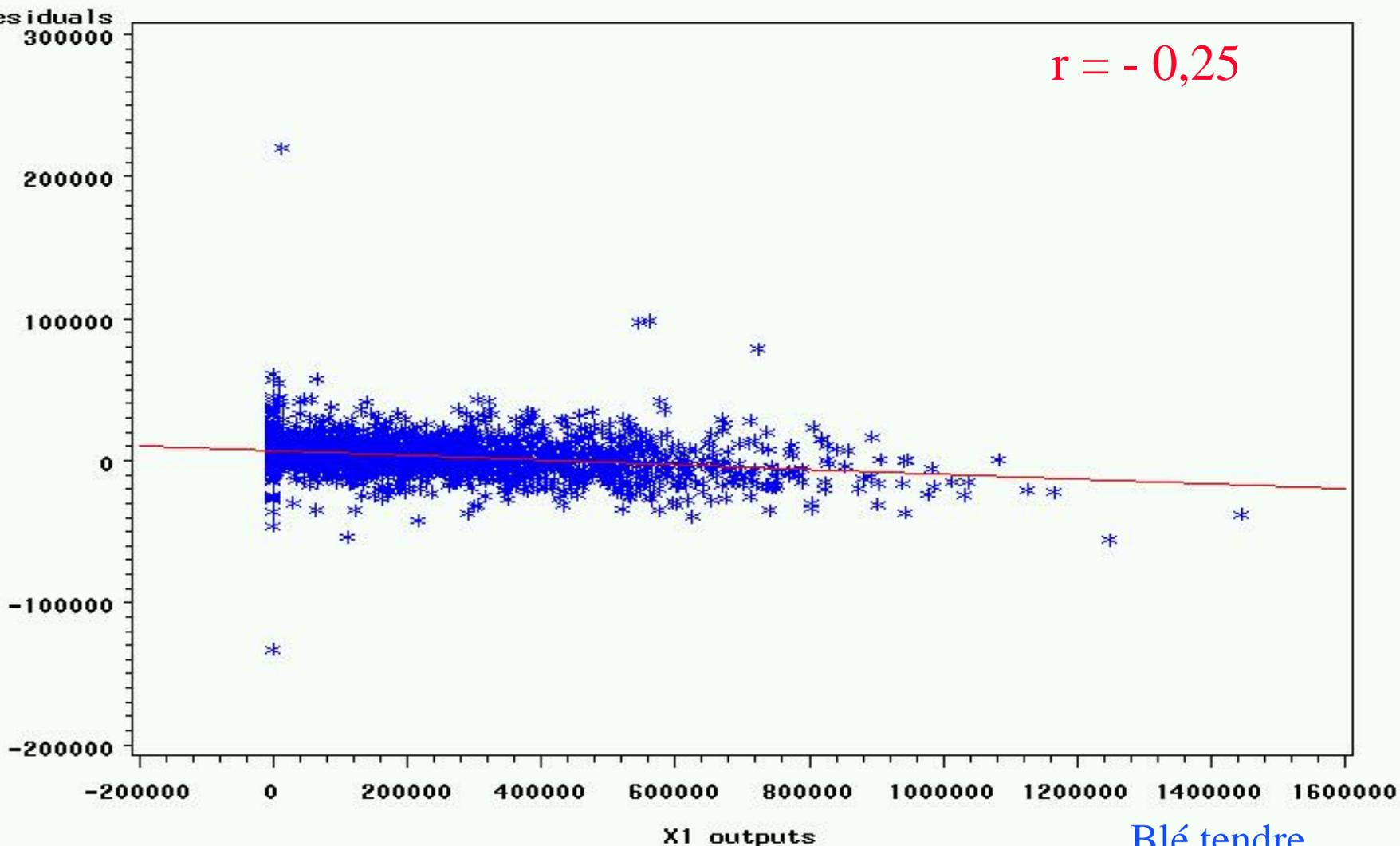
■ solutions:

- les méthodes d'estimation par variables instrumentales (doubles moindres carrés) peuvent être envisagées
- mais se heurtent à la difficulté de trouver des instruments adéquats



Assurances

RESIDUALS & OUTPUTS



Blé tendre

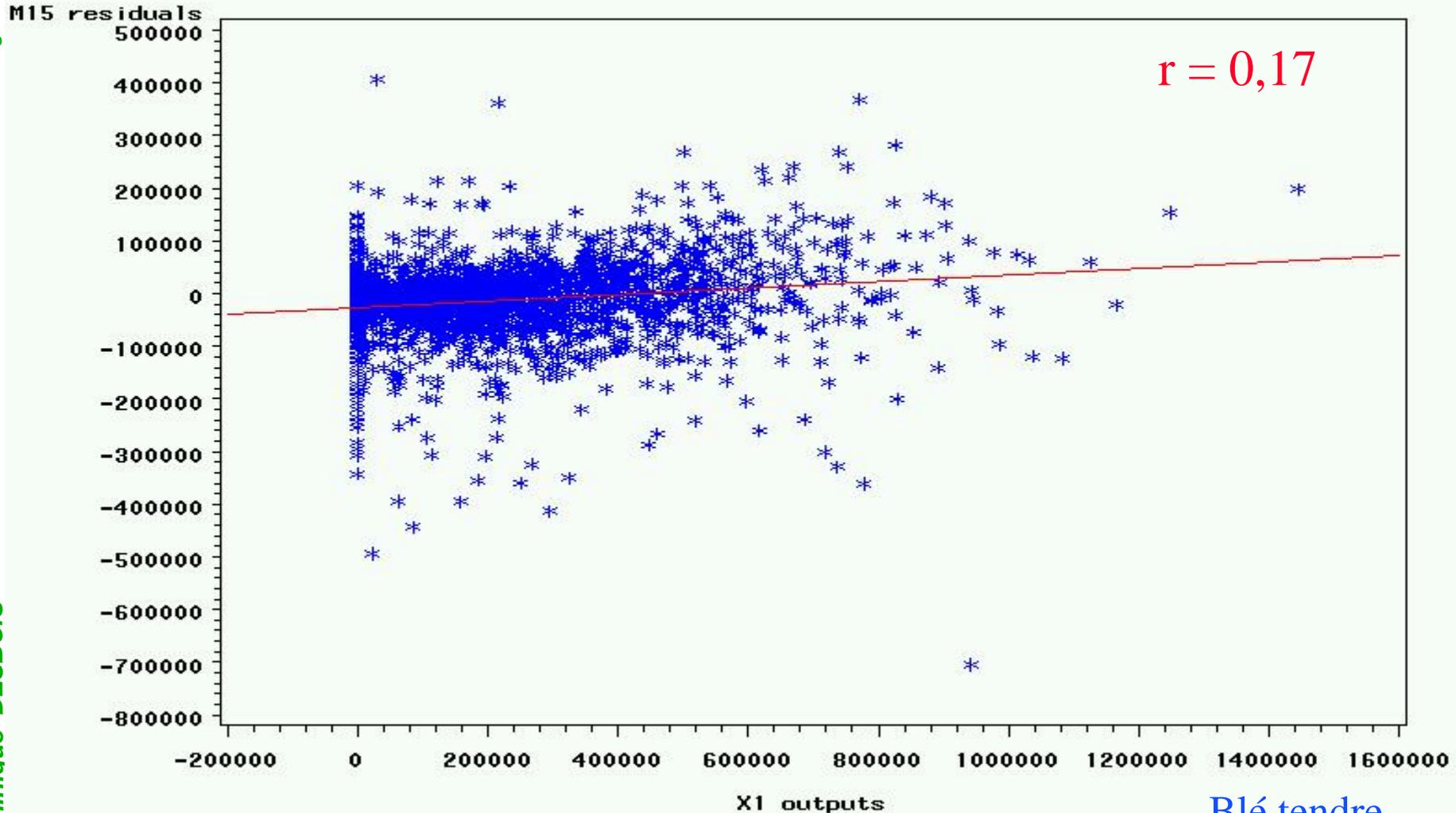
Econométrie des coûts de production agricole

Problèmes d'estimation : corrélations entre résidus et régresseurs



subventions d'exploitation

RESIDUALS & OUTPUTS





Problèmes d'estimation : pseudocolinéarité en régression multiple

- problèmes : *instabilité des coefficients de régression, difficulté d'interprétation de l'équation*
- détection :
 - *facteur d'inflation de la variance*
 - *indice de conditionnement maximal (BKW)*

$$VIF = \frac{1}{1 - R_j^2}$$

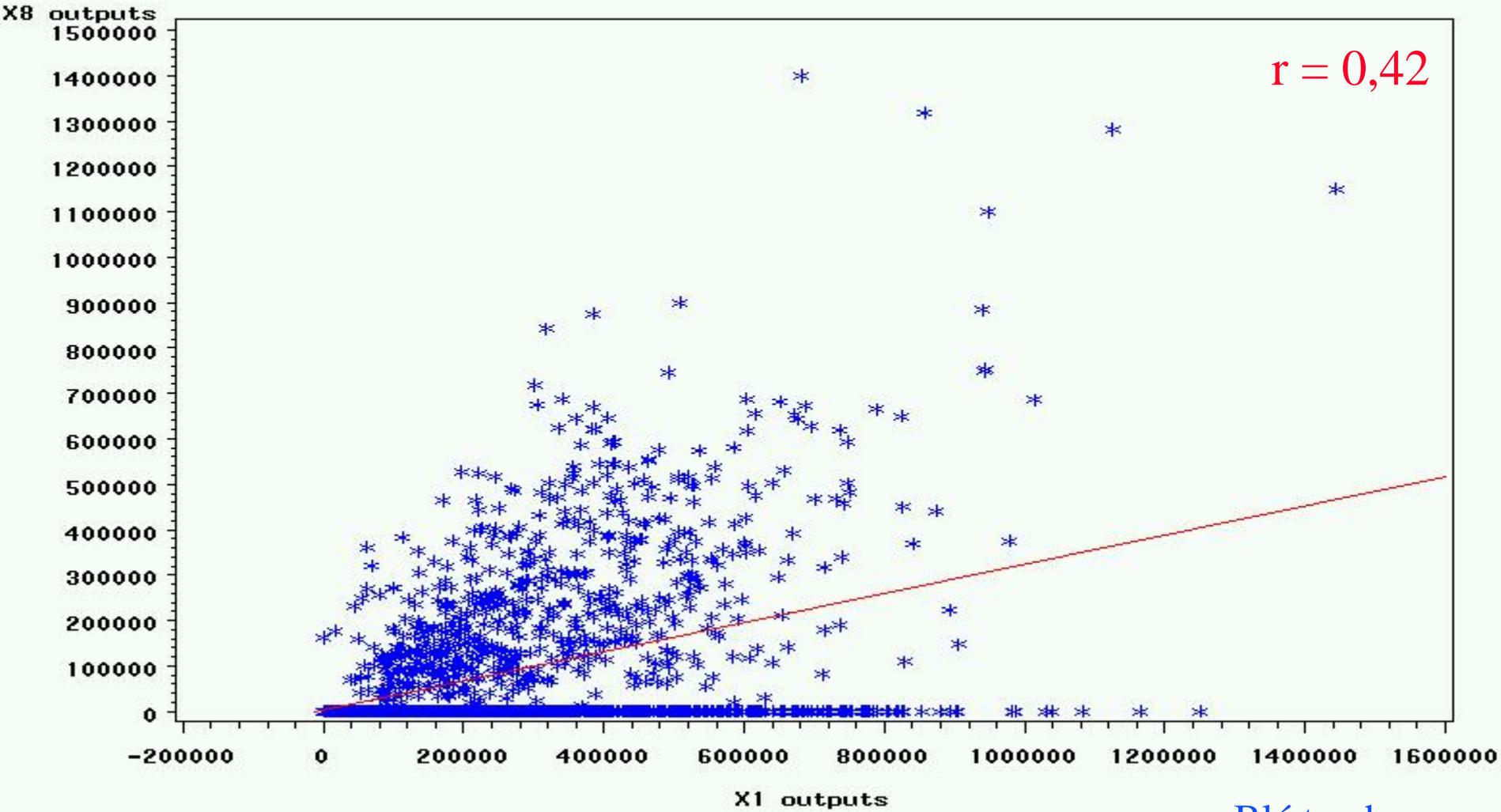
$$\mu_j = \sqrt{\frac{\lambda_1}{\lambda_{\min}}}$$

- solutions:
 - *régression pseudo-orthogonalisée*
 - *régression sur variables latentes*
 - *sélection de sous-ensembles de variables*



Betterave

OUTPUT COLINEARITY

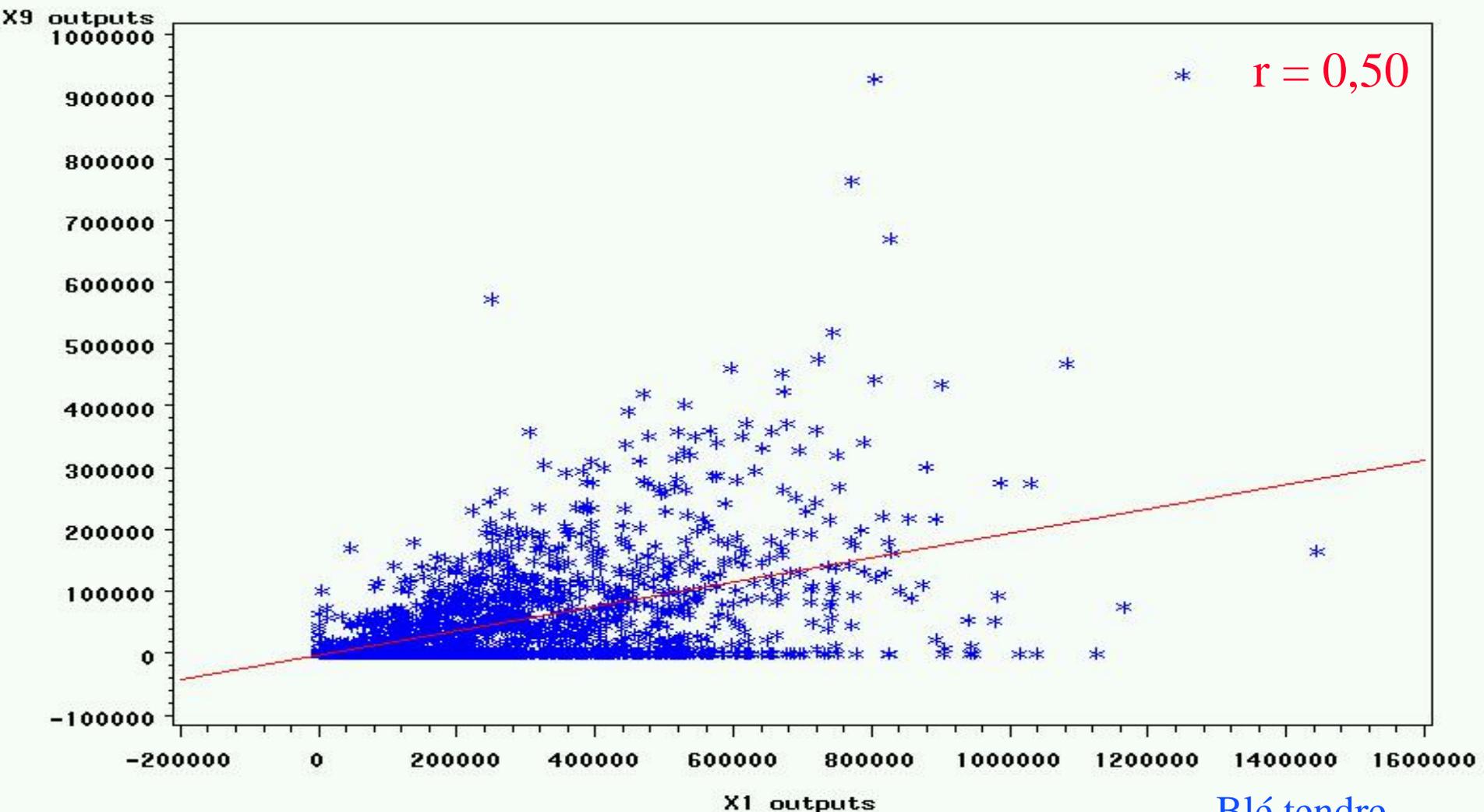


Blé tendre



Colza

OUTPUT COLLINEARITY

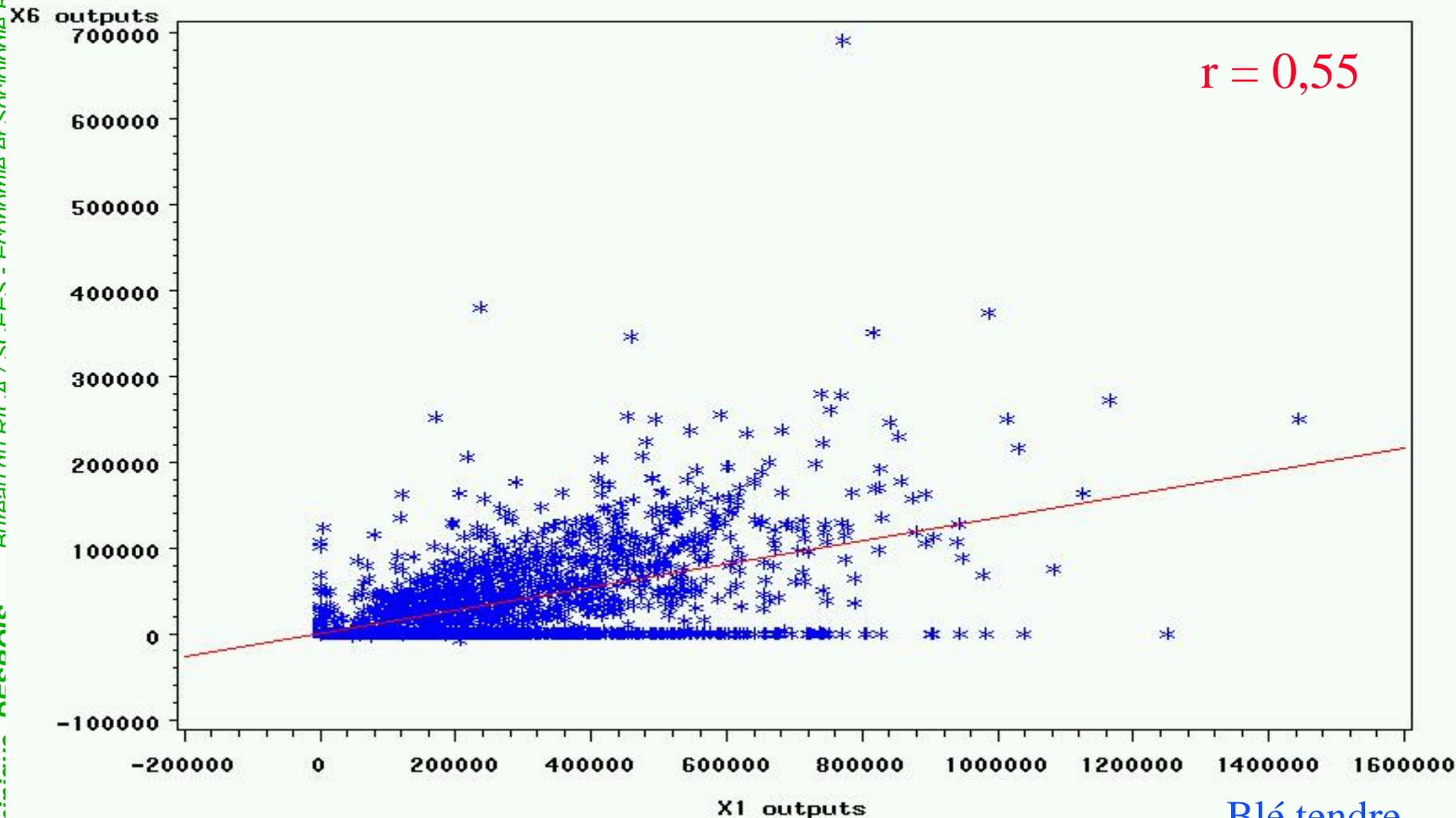


Blé tendre



Mais

OUTPUT COLINEARITY



Blé tendre



*Estimation d'un modèle à équations simultanées avec
contrainte sur les coefficients par*

*la **PROC SYSLIN** de SAS*

Procédure d'estimation :

SUR - "Seemingly Unrelated Regression"

$$Y_1 = \sum_{j=1}^p a_{j1} * X_j + U_1$$

$$Y_k = \sum_{j=1}^p a_{jk} * X_j + U_k$$

$$Y_q = \sum_{j=1}^p a_{jq} * X_j + U_q$$

Contraintes :

■ *la somme des coefficients pour chaque produit j
est égale à 1*

$$\sum_k a_{jk} + b_j = 1 \quad \forall j$$

■ *les coefficients de charges aliments et frais
spécifiques animaux k_a sont nuls pour l'ensemble
des produits végétaux j_v*

$$a_{j_v, k_a} = 0 \quad \forall (j_v, k_a)$$



Origine du modèle de régression sur variables latentes

- *économétrie* : *Hermann Wold*
 - *algorithme NILES*
 - *approche PLS*
- *chimie* : *Svante Wold*
 - *chromatographie*
 - *spectrographie*
- *régression sur variables latentes* : *Tenenhaus*
 - *La régression PLS, théorie et pratique, Technip, 1998.*



Le modèle de régression sur variables latentes

■ problème:

régresser Y (endogène) sur X (exogène) multicollinéaire

■ solution:

- *1) description de X par des composantes orthogonales*
- *2) régression de Y sur les composantes X les mieux corrélées à Y*
- *3) écriture de l'équation de Y sur X*

■ application:

- *régression sur composantes principales (RCP)*
- *régression sur les composantes des variables instrumentales (RACPVI)*
- *régression des moindres carrés partiels (RMCP - PLS)*

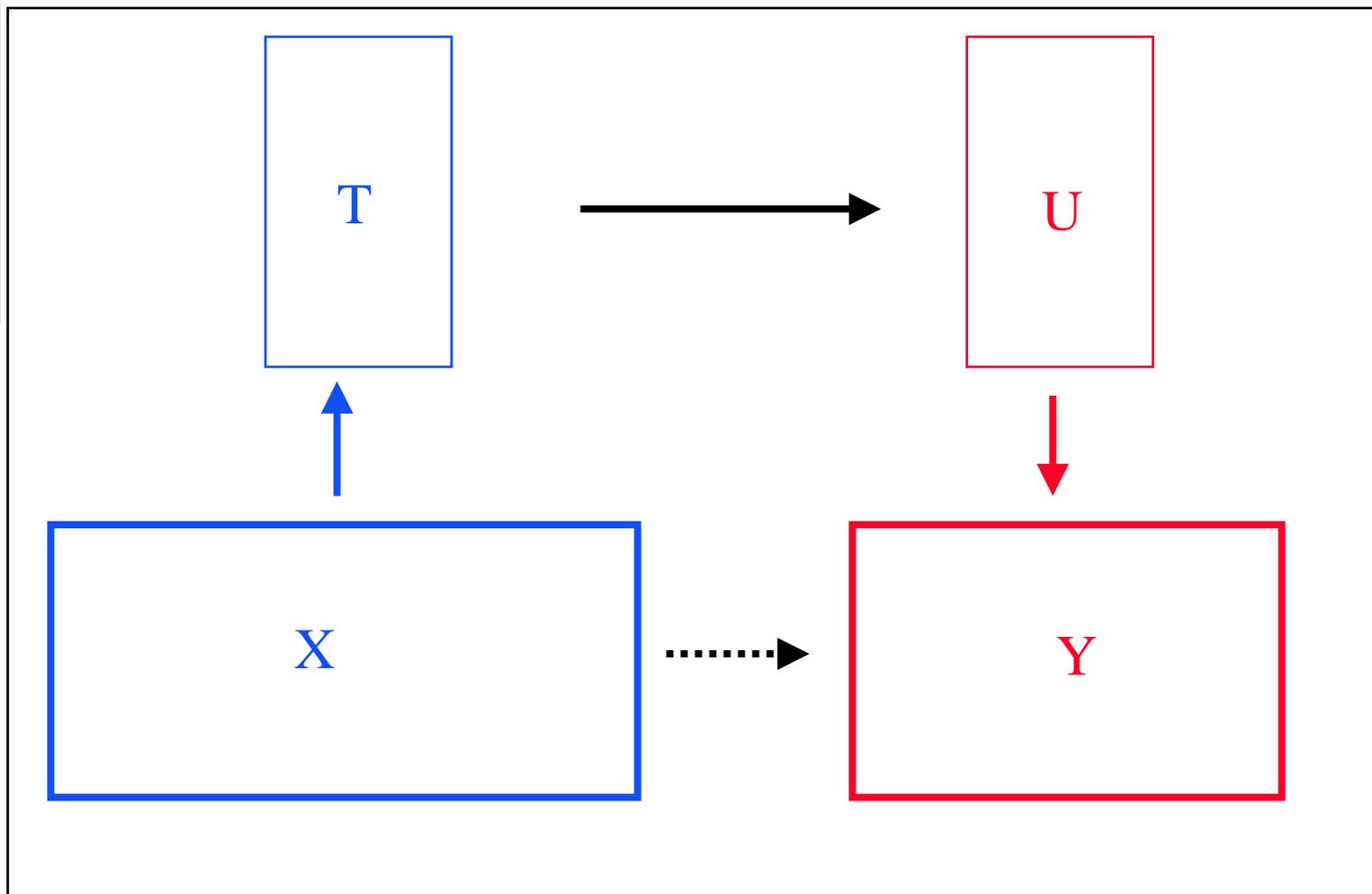


Objectif du modèle de régression des moindres carrés partiels

- extraire des pseudo-composantes orthogonales T à partir des variables explicatives X (facteurs exogènes)
- pour estimer à partir des pseudo-composantes orthogonales U qui leur sont liées les variables expliquées Y (réponses endogènes)



Modélisation sur variables latentes par la régression des moindres carrés partiels





La régression des moindres carrés partiels (MCP)

- *obtenir une liaison optimale au sein de chaque paire de X-pseudocomposantes T et de Y-pseudocomposantes U*
- *en recherchant les directions de l'espace des facteurs X liées aux plus fortes variations dans l'espace des réponses Y*
- *par décomposition en valeurs singulières de l'opérateur d'inertie $X'Y$*



Le modèle théorique de la régression MCP

■ $X = TP' + E$

- X : matrice des facteurs
- T : matrice des X -pseudocomposantes
- P : matrice des X -saturations
- E : matrice des X -résidus

■ $Y = UQ' + F$

- Y : matrice des réponses
- U : matrice des Y -pseudocomposantes
- Q : matrice des Y -saturations
- F : matrice des Y -résidus



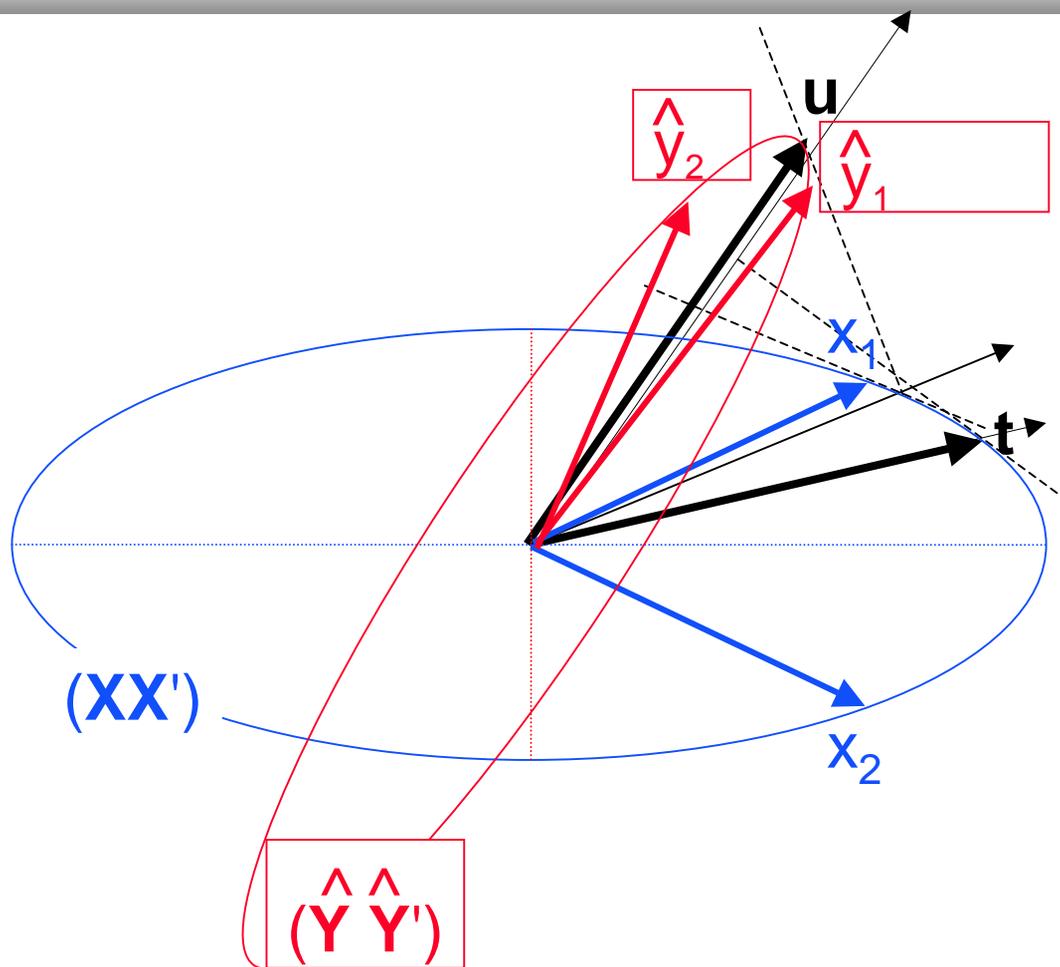
Les algorithmes de la régression des moindres carrés partiels (MCP)

- *NIPALS 1 : ACP sans données manquantes*
- *NIPALS 2 : estimation d'un modèle bilinéaire avec données manquantes*
- *PLS 1 : régression MCP pour y univarié*
- *PLS 2 : régression MCP pour Y multivarié*
- *SIMPLS (et sa variante SIMPLS-CA)*



Algorithme PLS2

- 1** "Initialisation": $X_0 = X$; $Y_0 = Y$;
- 2** Pour $k=1, 2, \dots, K$ faire { début
 - 2.1** $u_k = Y_{k-1} [\cdot ; 1]$;
 - 2.2** Faire jusqu'à $|w_k - w_k^*| < \varepsilon$ { début
 - 2.2.1** $w_k = X'_{k-1} u_k / u'_k u_k$; * projection de X sur u
 - 2.2.2** $w_k = w_k / |w_k|$; * normalisation de w
 - 2.2.3** $t_k = X_{k-1} w_k / w'_k w_k$; * calcul de t
 - 2.2.4** $c_k = Y'_{k-1} t_k / t'_k t_k$; * projection Y sur t
 - 2.2.5** $u_k = Y_{k-1} c_k / c'_k c_k$; * calcul de u
 - 2.3** $p_k = X'_{k-1} t_k / t'_k t_k$; * projection de X sur t
 - 2.4** $X_k = X_{k-1} - t_k p'_k$; * mises a jour de X et de Y
 - 2.5** $Y_k = Y_{k-1} - t_k c'_k$; * fin }





Régression PLS2

Validation croisée : le critère PRESS

- *modèle à*
 - L* *pseudo-composantes,*
 - K* *variables expliquées*
 - J* *variables explicatives,*
 - I* *observations*

Somme des carrés des erreurs d'estimations

$$PRESS_l = \sum_i (y_i - \hat{y}_{(-i)}^l)^2$$

Règle empirique

la l-ième composante t_l est retenue si :

$$\sqrt{PRESS_l} \leq 0,95 * \sqrt{PRESS_{l-1}}$$



- *différences entre les carrés des erreurs d'estimations*

$$D_{l,ik} = R_{l,ik}^2 - R_{l_{\min},ik}^2$$

- *valeur critique*

$$C_l = \sum_{ik} D_{l,ik}$$

- *T² de Hotelling*

$$C_l = d_l' S_l d_l$$

$$S_l = \sum_i d_{l,i} d_{l,i}'$$

$$d_l = \sum_i d_{l,i}$$

$$d_{l,i} = \{D_{l,i1}, \dots, D_{l,iK}\}$$



Estimation PLS2 :

validation croisée avec le critère PRESS

Random Subset Validation for the Number of Extracted Factors

The PLS Procedure

Number of Extracted Factors	Root Mean PRESS	Prob > PRESS
0	1.155714	<.0001
1	1.067863	<.0001
2	1.008234	<.0001
3	0.938102	<.0001
4	0.920035	<.0001
5	0.919288	<.0001
6	0.923448	<.0001
7	0.903448	<.0001
8	0.890096	<.0001
9	0.889128	<.0001
10	0.885717	<.0001
11	0.891929	<.0001
12	0.872239	0.0550
13	0.876255	0.4340
14	0.87337	0.2300
15	0.871804	1.0000

Minimum root mean PRESS 0.8718

Minimizing number of factors 15

Smallest number of factors with $p > 0.1$ 13

$$\sqrt{0.876255} \leq 0.95 * \sqrt{0.872239} \approx 0.887240$$



Régression PLS2

modèle à 13 pseudocomposantes : part de variance expliquée

REGRESSION PLS - PROC PLS SAS
 modèle à 13 pseudo-composantes
 RICA & MODELE COUT DE PRODUCTION

The PLS Procedure
 Percent Variation Accounted for
 by Partial Least Squares Factors

Number of Extracted Factors	Model Effects		Dependent Variables	
	Current	Total	Current	Total
1	11.0772	11.0772	23.3691	23.3691
2	3.8535	14.9307	20.6884	44.0575
3	5.3192	20.2499	8.8924	52.9499
4	4.2710	24.5209	4.4220	57.3719
5	4.4904	29.0113	3.0928	60.4646
6	4.3857	33.3970	1.7988	62.2634
7	4.1595	37.5564	1.0815	63.3449
8	4.9310	42.4874	0.7396	64.0844
9	4.2160	46.7034	0.4941	64.5786
10	4.0266	50.7300	0.4317	65.0103
11	3.8846	54.6146	0.2609	65.2712
12	3.9523	58.5669	0.1890	65.4603
13	4.0218	62.5888	0.1395	65.5998



*Statistique du R^2 matriciel pondéré
et de l'écart moyen matriciel pondéré*

The SYSLIN Procedure

Seemingly Unrelated Regression Estimation

System Weighted MSE	0.9373
Degrees of freedom	121337
System Weighted R-Square	0.7927

$$R^2 = \frac{Y'WR(X'X)^{-1}R'WY}{Y'WY}$$

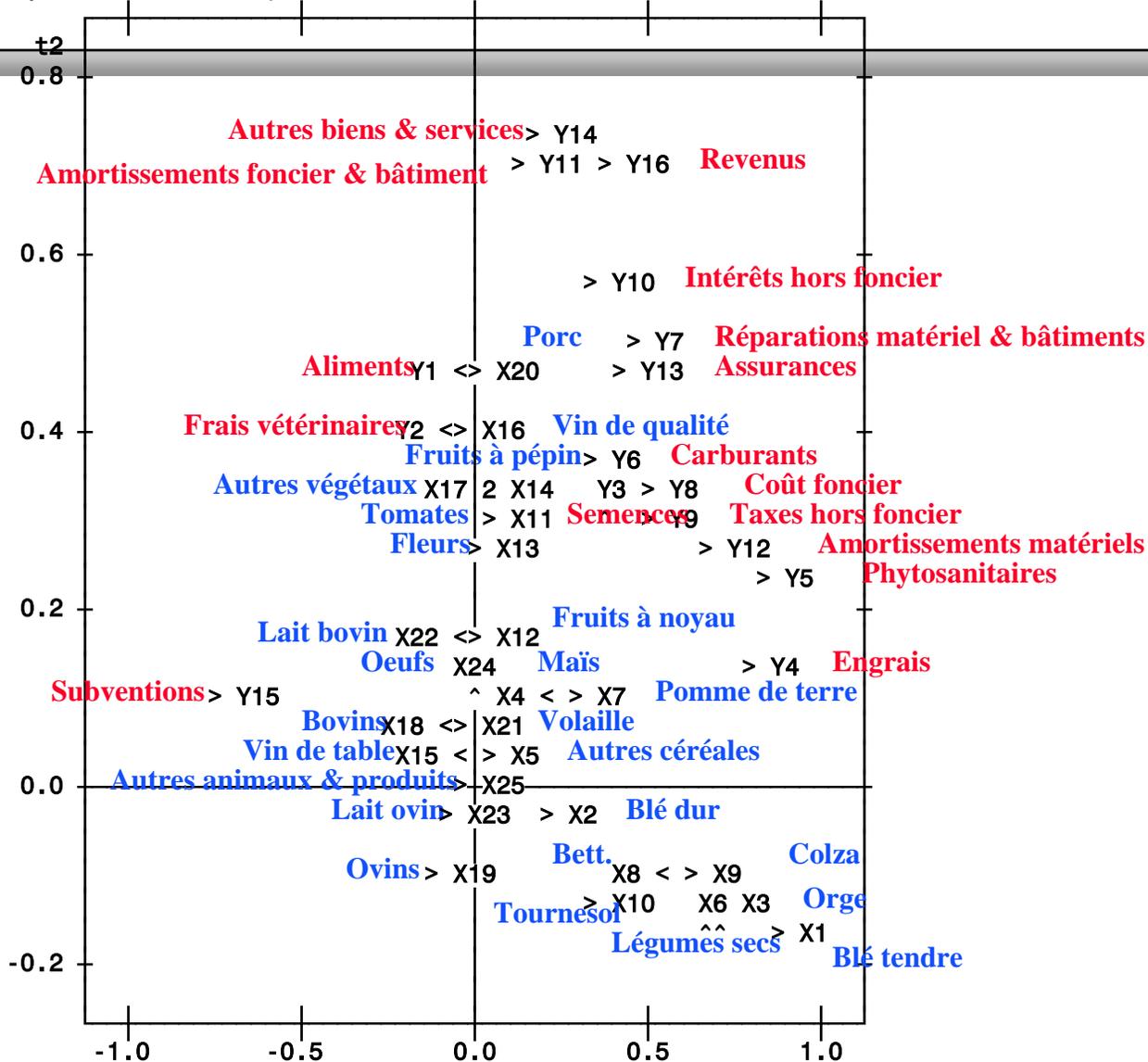
$$X'X = R'WR$$

$$W = S^{-1} \otimes Z(Z'Z)^{-1}Z'$$

$$MSE = \frac{1}{tdf} \left(Y'WY - Y'WR(X'X)^{-1}R'WY \right)$$

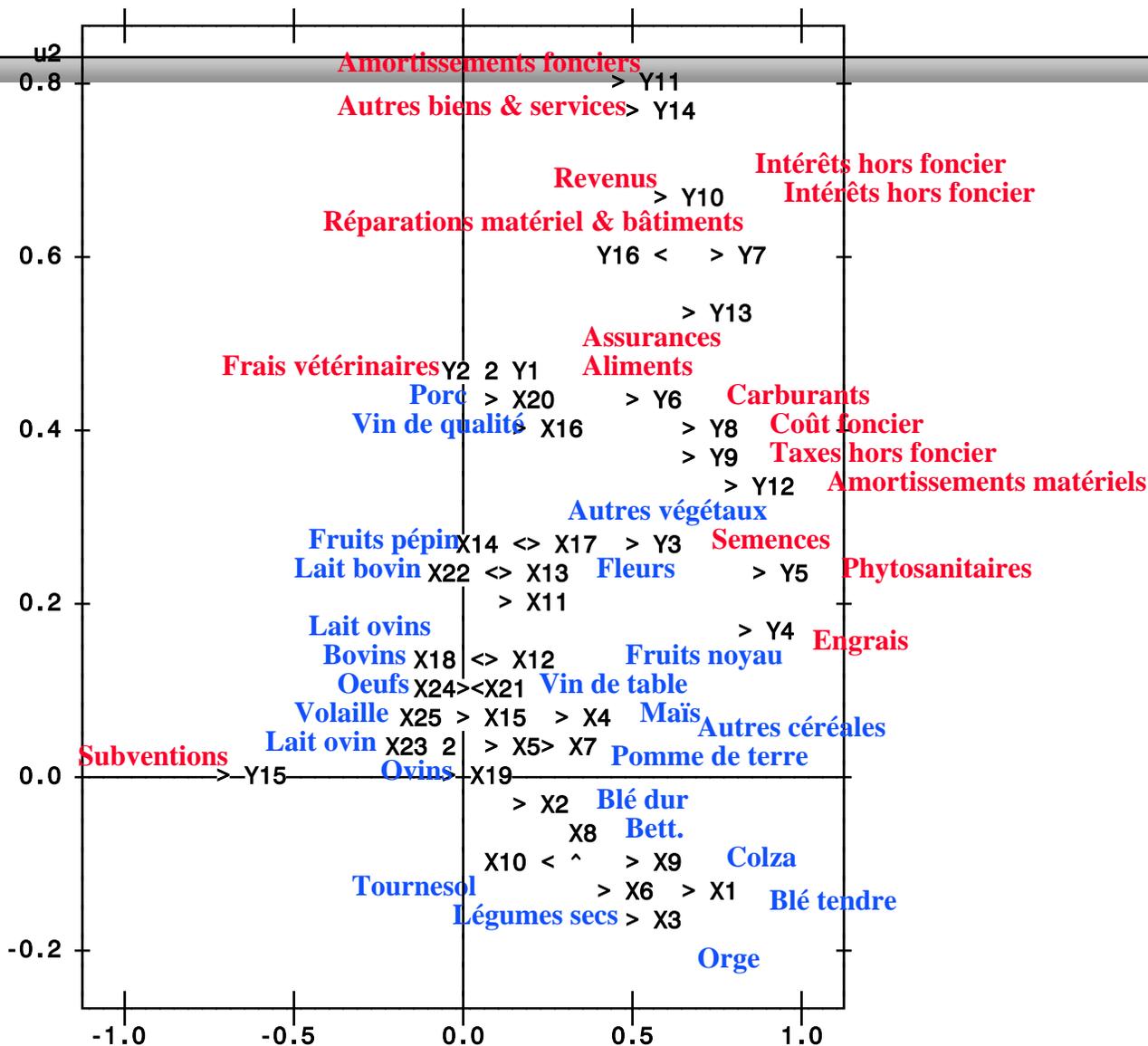


Graphique des variables, corrélations avec t1 x t2



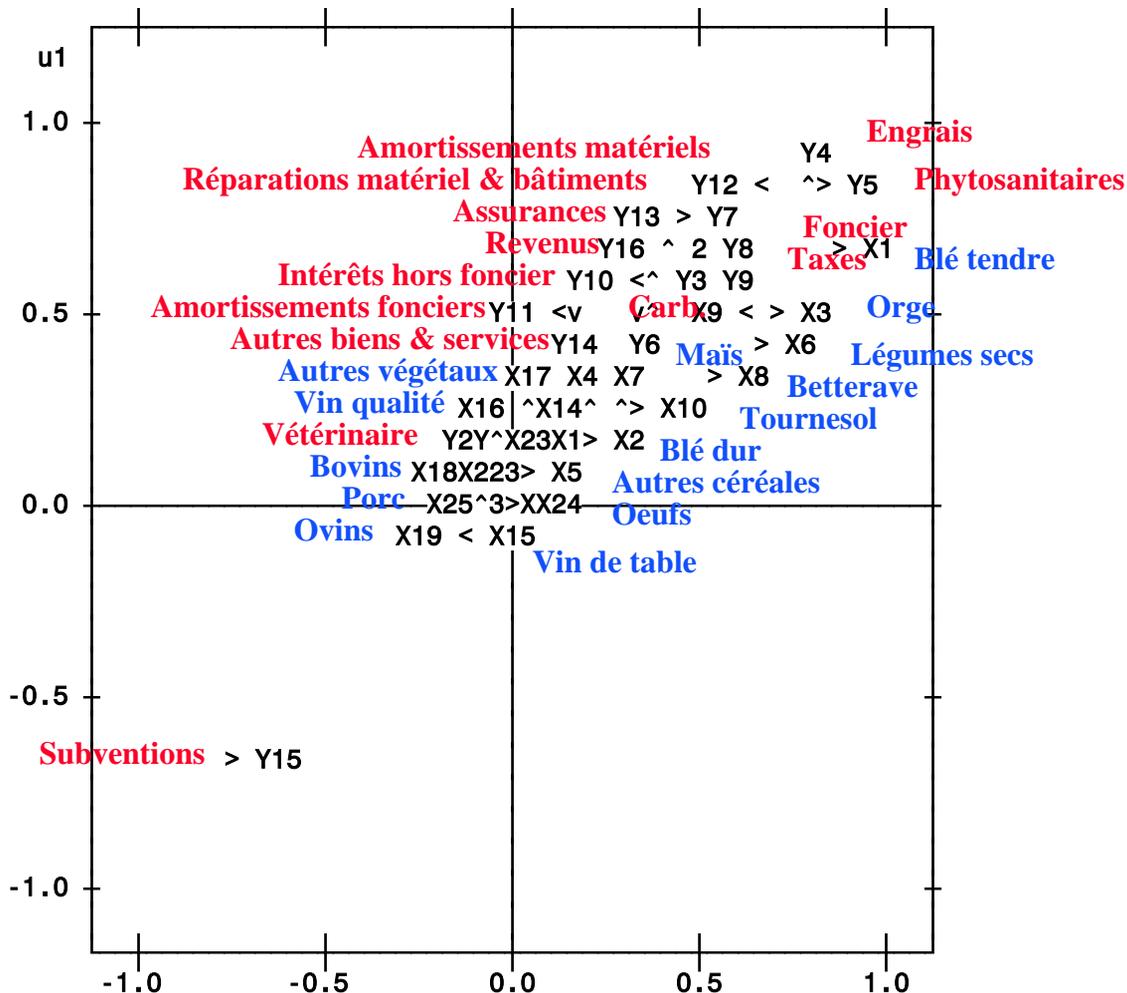


Graphique des variables, corrélations avec u1 x u2





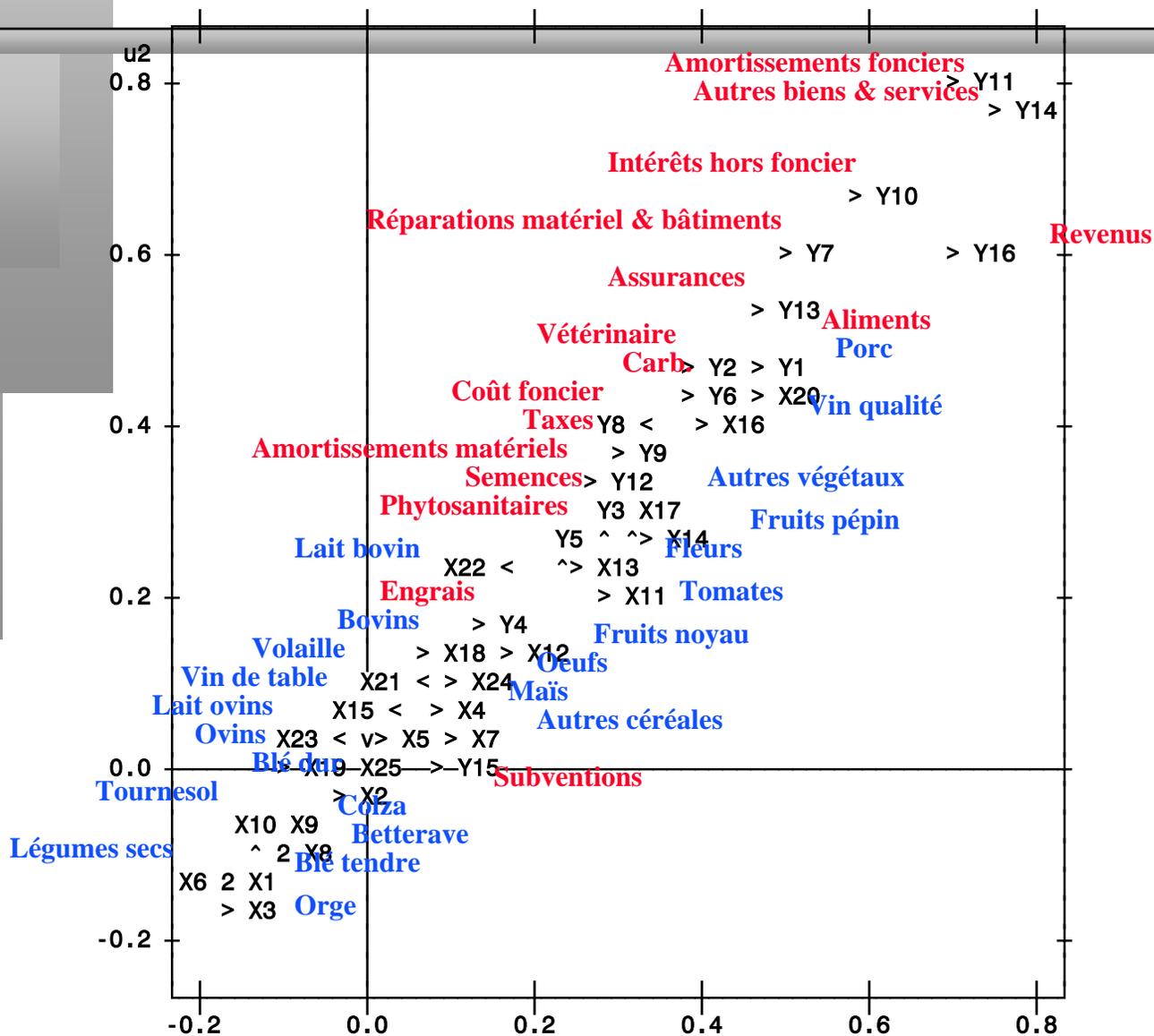
Graphique des variables, corrélations avec t1 x u1





Ajustement entre secondes X- et Y-pseudocomposantes

Graphique des variables, corrélations avec $t_2 \times u_2$



*Econométrie des coûts de production agricole
Comparaison entre estimation PLS2 et estimation SUR :
corrélation entre charge estimée et charge observée*



corrélation de Yk et Yhk	SUR	PLS
Y1 ALIMENTS	0,96008	0,9595
Y2 VETERINAIRE	0,87100	0,86986
Y3 SEMENCES	0,79233	0,79062
Y4 ENGRAIS	0,87901	0,87549
Y5 PRODUITS PHYTOSANITAIRES	0,90642	0,90382
Y6 PRODUITS PETROLIERS	0,7447	0,74407
Y7 REPARATIONS BAT. & MAT.	0,7218	0,72094
Y8 COUT FONCIER	0,79558	0,79358
Y9 TAXES (HORS FONCIER)	0,64249	0,63701
Y10 INTERETS (HORS FONCIER)	0,72734	0,72613
Y11 AMORTISSEMENTS FONC BAT	0,73499	0,7334
Y12 AMORTISSEMENTS MAT	0,77919	0,77734
Y13 ASSURANCES	0,73564	0,74685
Y14 AUTRES BIENS ET SERVICES	0,83061	0,82941
Y15 AUTRES SUBVENTIONS	0,91127	0,91144
Y16 REVENU	0,87257	0,86917

*Econométrie des coûts de production agricole
Comparaison entre estimation PLS2 et estimation SUR :
corrélation entre charge estimée et charge observée*



corrélation de Yk et Yhk	SUR	PLS
Y1 ALIMENTS	0,96008	0,9595
Y2 VETERINAIRE	0,87100	0,86986
Y3 SEMENCES	0,79233	0,79062
Y4 ENGRAIS	0,87901	0,87549
Y5 PRODUITS PHYTOSANITAIRES	0,90642	0,90382
Y6 PRODUITS PETROLIERS	0,7447	0,74407
Y7 REPARATIONS BAT. & MAT.	0,7218	0,72094
Y8 COUT FONCIER	0,79558	0,79358
Y9 TAXES (HORS FONCIER)	0,64249	0,63701
Y10 INTERETS (HORS FONCIER)	0,72734	0,72613
Y11 AMORTISSEMENTS FONC BAT	0,73499	0,7334
Y12 AMORTISSEMENTS MAT	0,77919	0,77734
Y13 ASSURANCES	0,73564	0,74685
Y14 AUTRES BIENS ET SERVICES	0,83061	0,82941
Y15 AUTRES SUBVENTIONS	0,91127	0,91144
Y16 REVENU	0,87257	0,86917

Econométrie des coûts de production agricole

Comparaison entre estimation PLS2 et estimation SUR : écart-moyen quadratique



<i>Estimations</i>	<i>SUR</i>	<i>PLS</i>
<i>écart-moyen quadratique</i>		
Y1 ALIMENTS	63513	65679
Y2 VETERINAIRE	14313	14703
Y3 SEMENCES	40829	44151
Y4 ENGRAIS	27994	31503
Y5 PRODUITS PHYTOSANITAIRES	25868	30607
Y6 PRODUITS PETROLIERS	20440	21034
Y7 REPARATIONS BAT. & MAT.	29000	31011
Y8 COUT FONCIER	46094	47989
Y9 TAXES (HORS FONCIER)	9789	10093
Y10 INTERETS (HORS FONCIER)	30868	32035
Y11 AMORTISSEMENTS FONC BAT	36628	39285
Y12 AMORTISSEMENTS MAT	52011	54608
Y13 ASSURANCES	15309	14802
Y14 AUTRES BIENS ET SERVICES	100066	119272
Y15 AUTRES SUBVENTIONS	66557	64330
Y16 REVENU	156730	158428
<i>System weighted MSE</i>	<i>0,9373</i>	
<i>Minimum root Mean Press</i>		<i>0,8718</i>

Econométrie des coûts de production agricole
Comparaison entre estimation PLS2 et estimation SUR :
R² ajusté

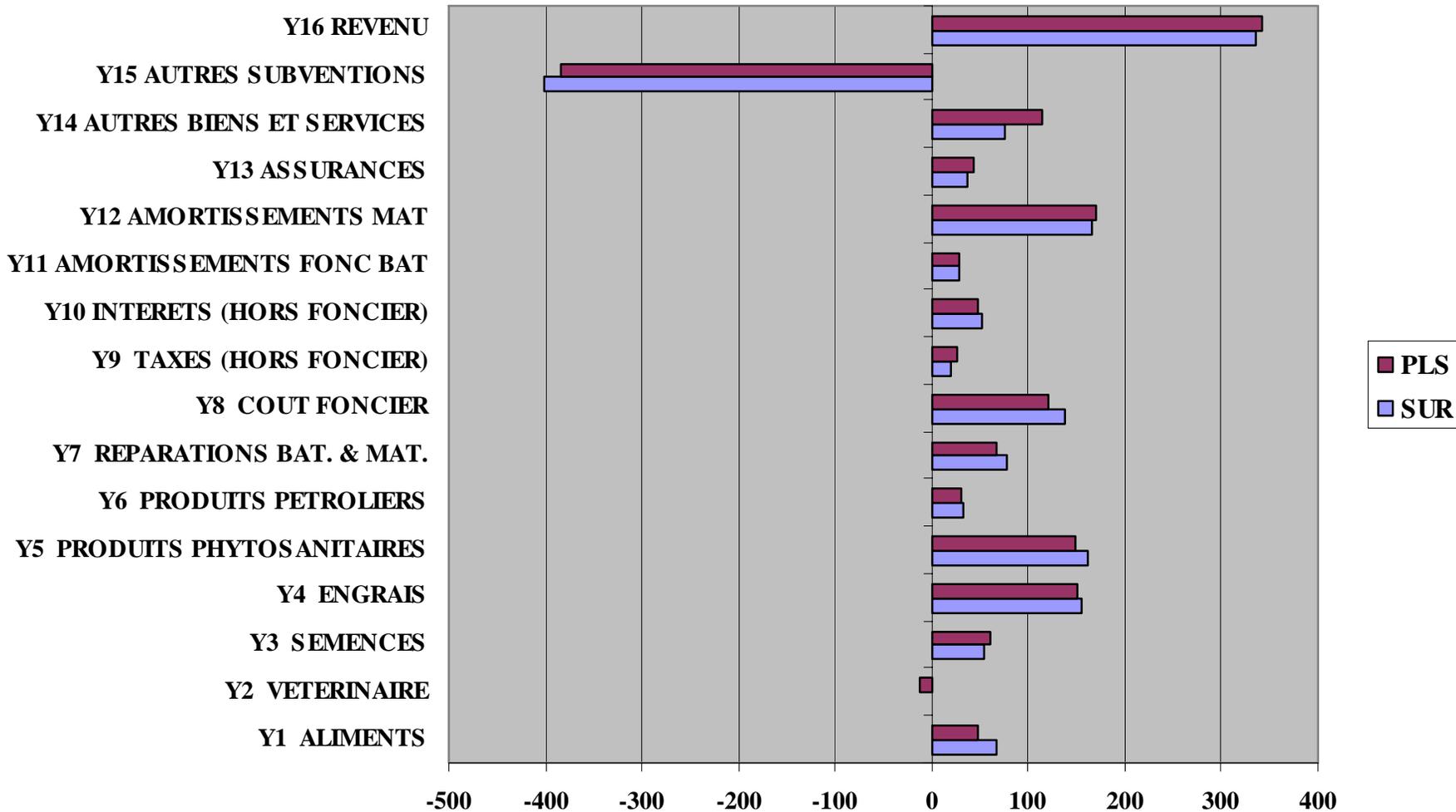


<i>Estimations</i>	<i>SUR</i>	<i>PLS</i>
<i>R2 ajusté de Yk</i>	<i>avec x1, ..., x25</i>	<i>avec t1, ..., t13</i>
Y1 ALIMENTS	0,93105	0,9159
Y2 VETERINAIRE	0,81697	0,7441
Y3 SEMENCES	0,70074	0,5634
Y4 ENGRAIS	0,88111	0,7079
Y5 PRODUITS PHYTOSANITAIRES	0,89169	0,7468
Y6 PRODUITS PETROLIERS	0,67552	0,5257
Y7 REPARATIONS BAT. & MAT.	0,76876	0,4404
Y8 COUT FONCIER	0,78058	0,6003
Y9 TAXES (HORS FONCIER)	0,63258	0,3708
Y10 INTERETS (HORS FONCIER)	0,68788	0,4911
Y11 AMORTISSEMENTS FONC BAT	0,69372	0,4671
Y12 AMORTISSEMENTS MAT	0,80258	0,5608
Y13 ASSURANCES	0,77778	0,5179
Y14 AUTRES BIENS ET SERVICES	0,79266	0,558
Y15 AUTRES SUBVENTIONS	0,90586	0,8331
Y16 REVENU	0,854	0,755
<i>System Weighted R-Square</i>	<i>0,7927</i>	
<i>Percent Variation Accounted</i>		<i>0,656</i>

Comparaison des résultats : le blé tendre

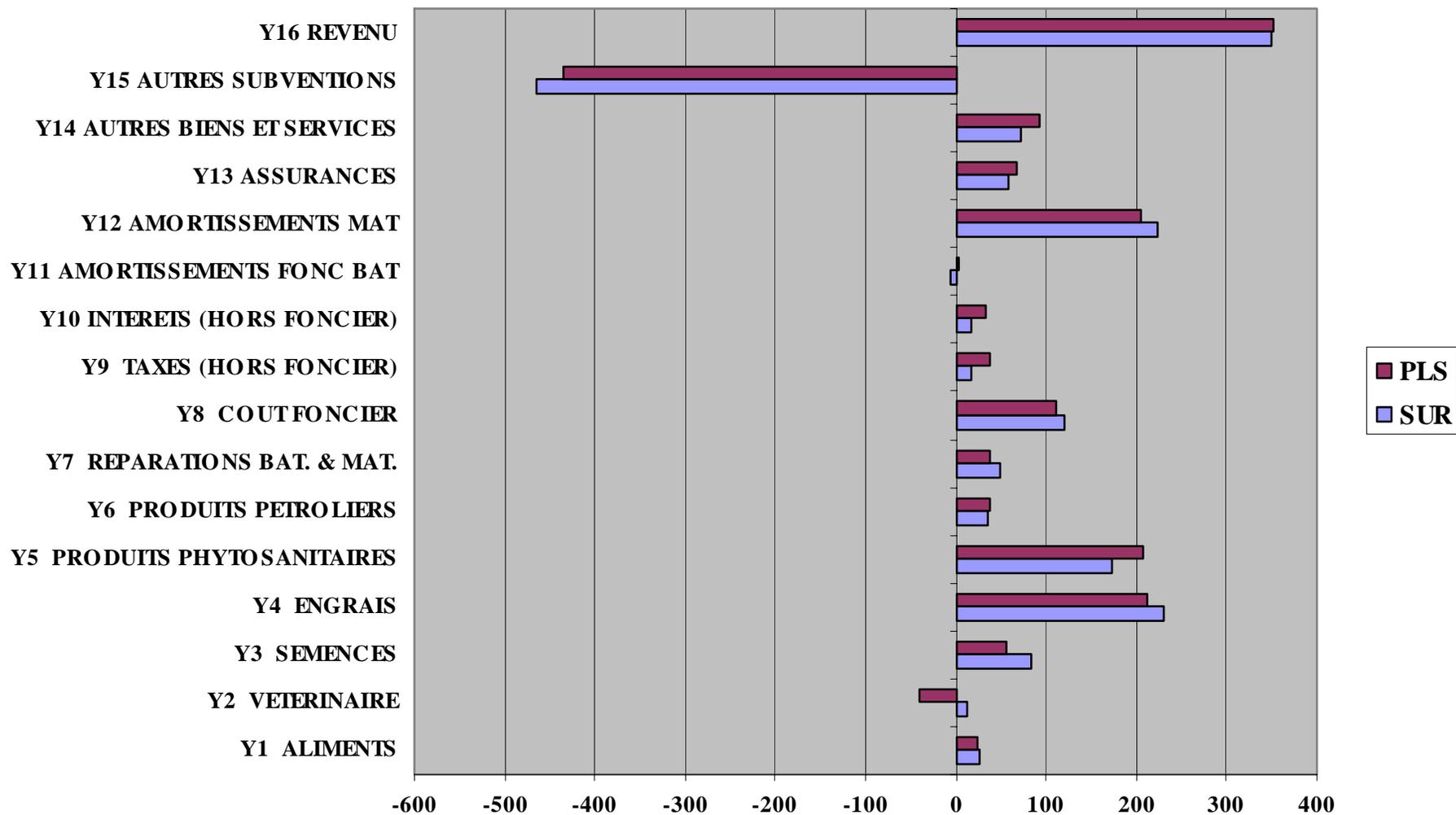


X1 blé tendre





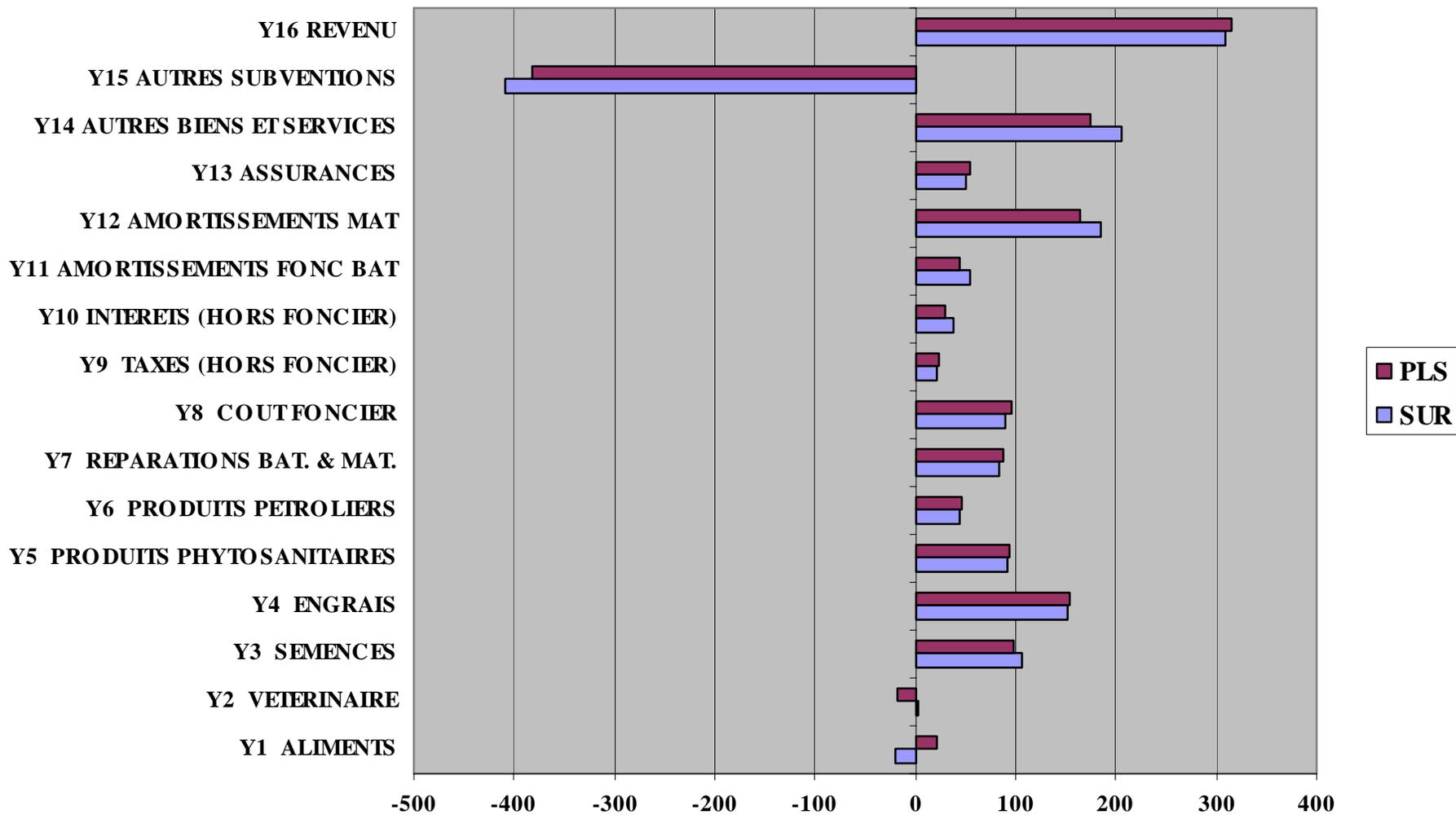
X3 orge





Comparaison des résultats : le maïs

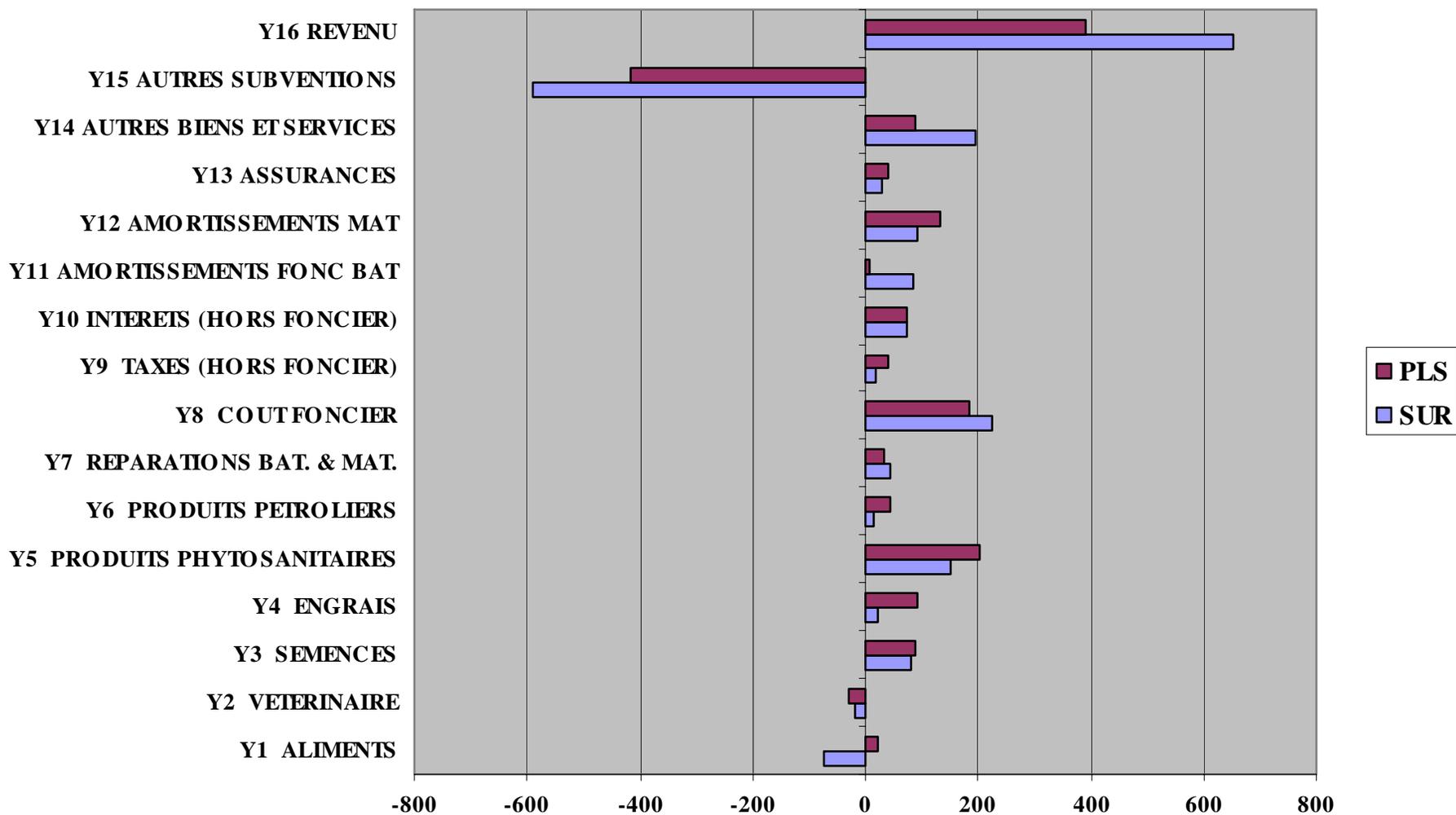
X4 maïs





Comparaison des résultats : les légumes secs

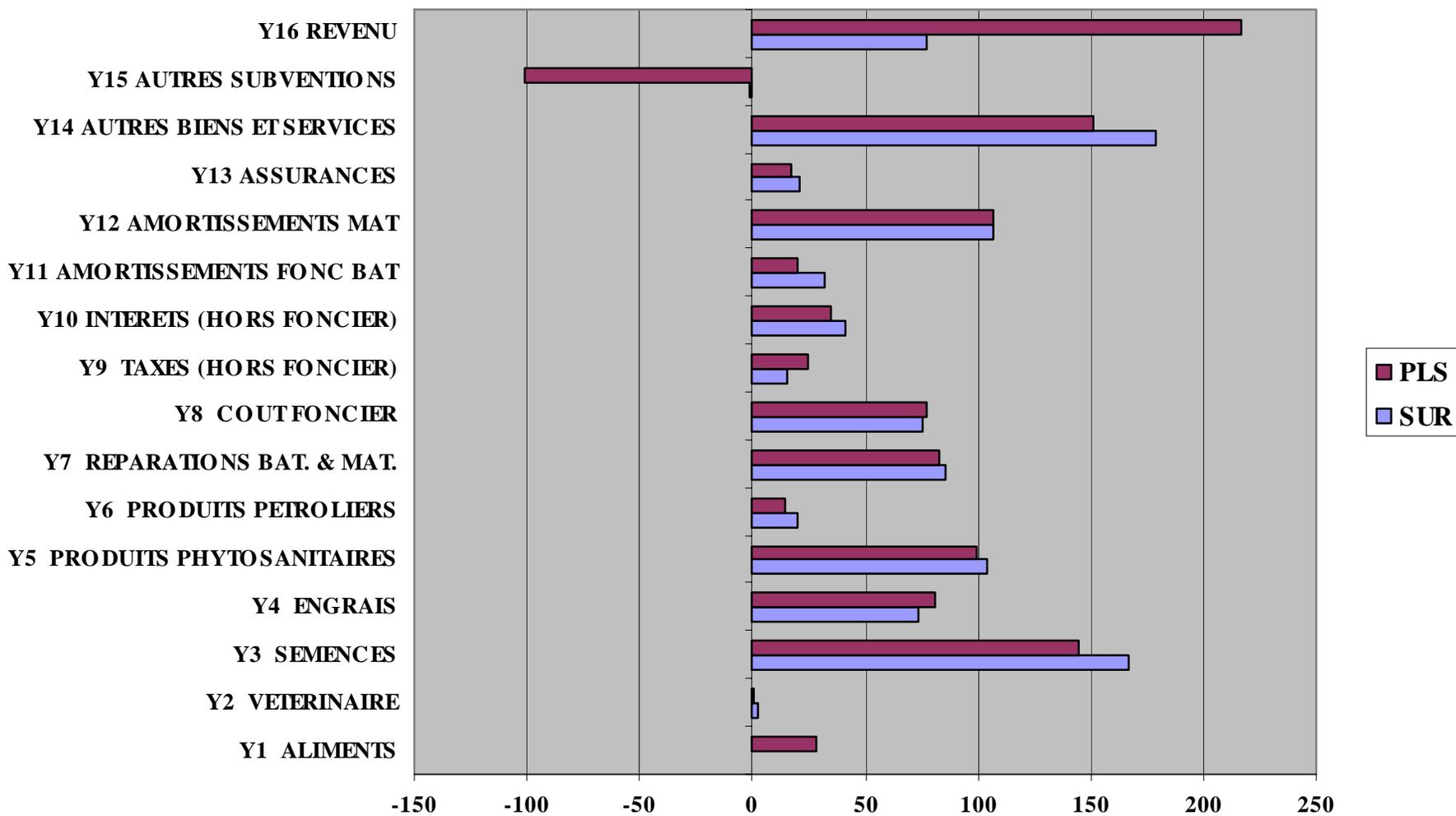
X6 légumes secs





Comparaison des résultats : la pomme de terre

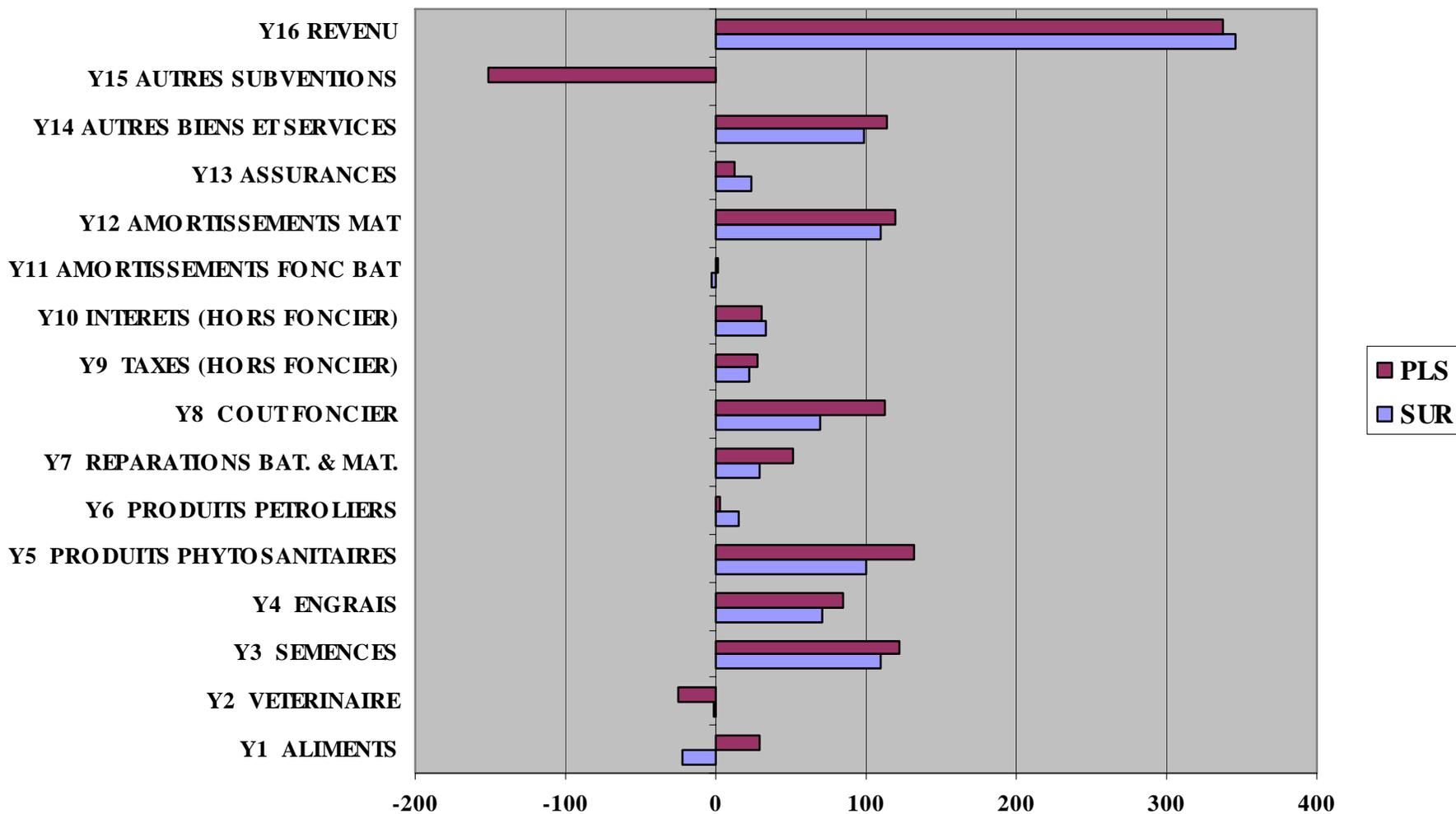
X7 pomme de terre





Comparaison des résultats : la betterave

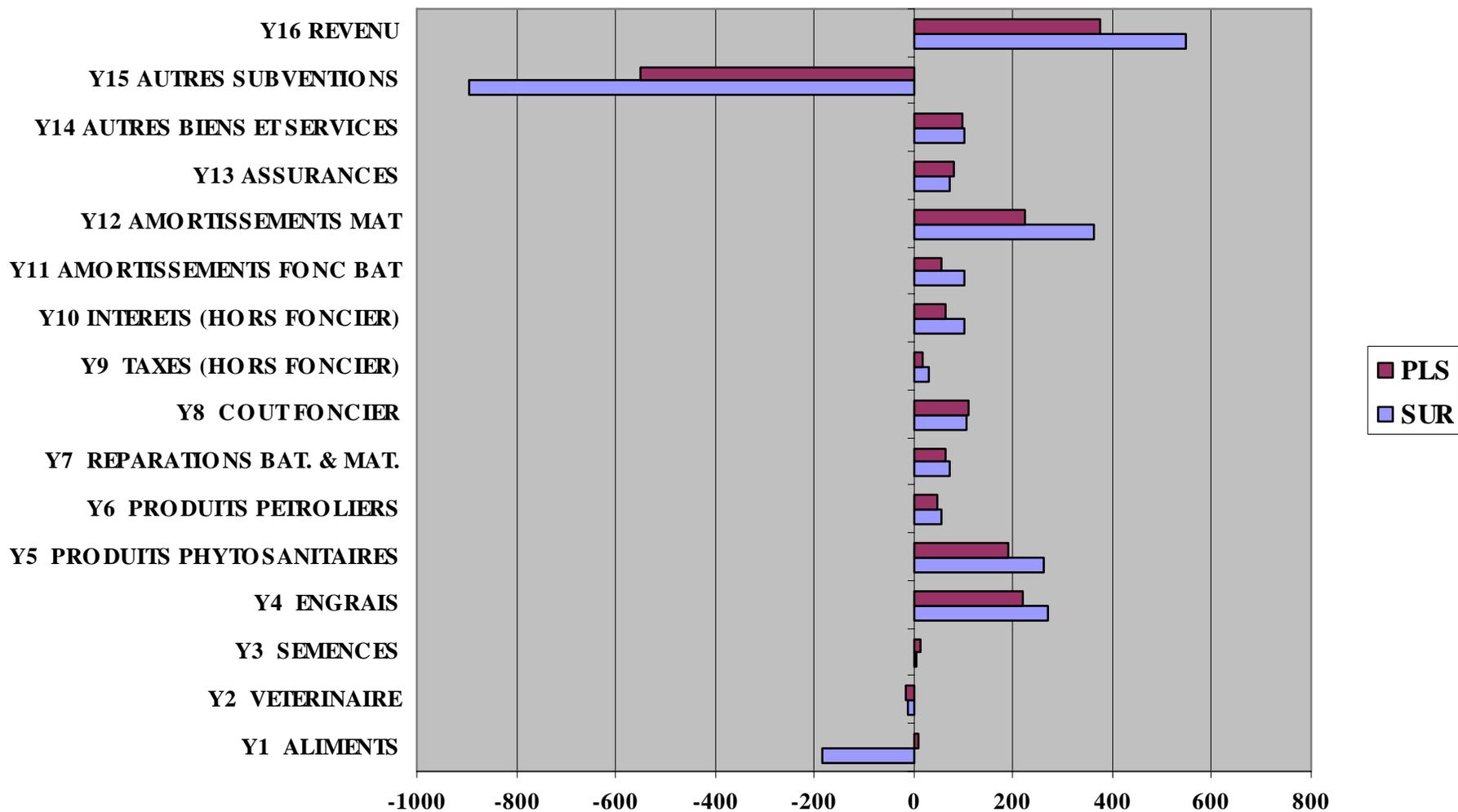
X8 betterave





Comparaison des résultats : le colza

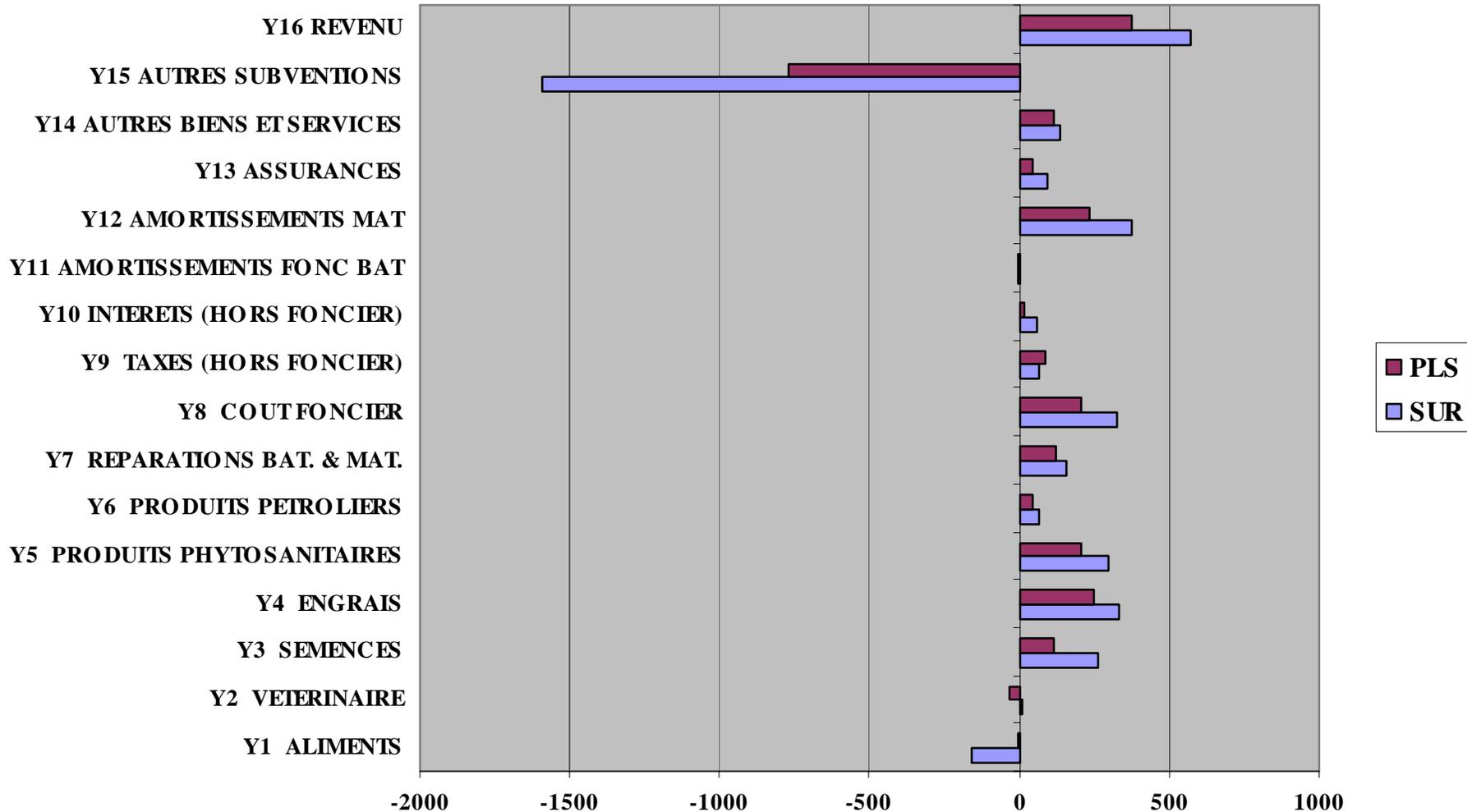
X9 colza





Comparaison des résultats : le tournesol

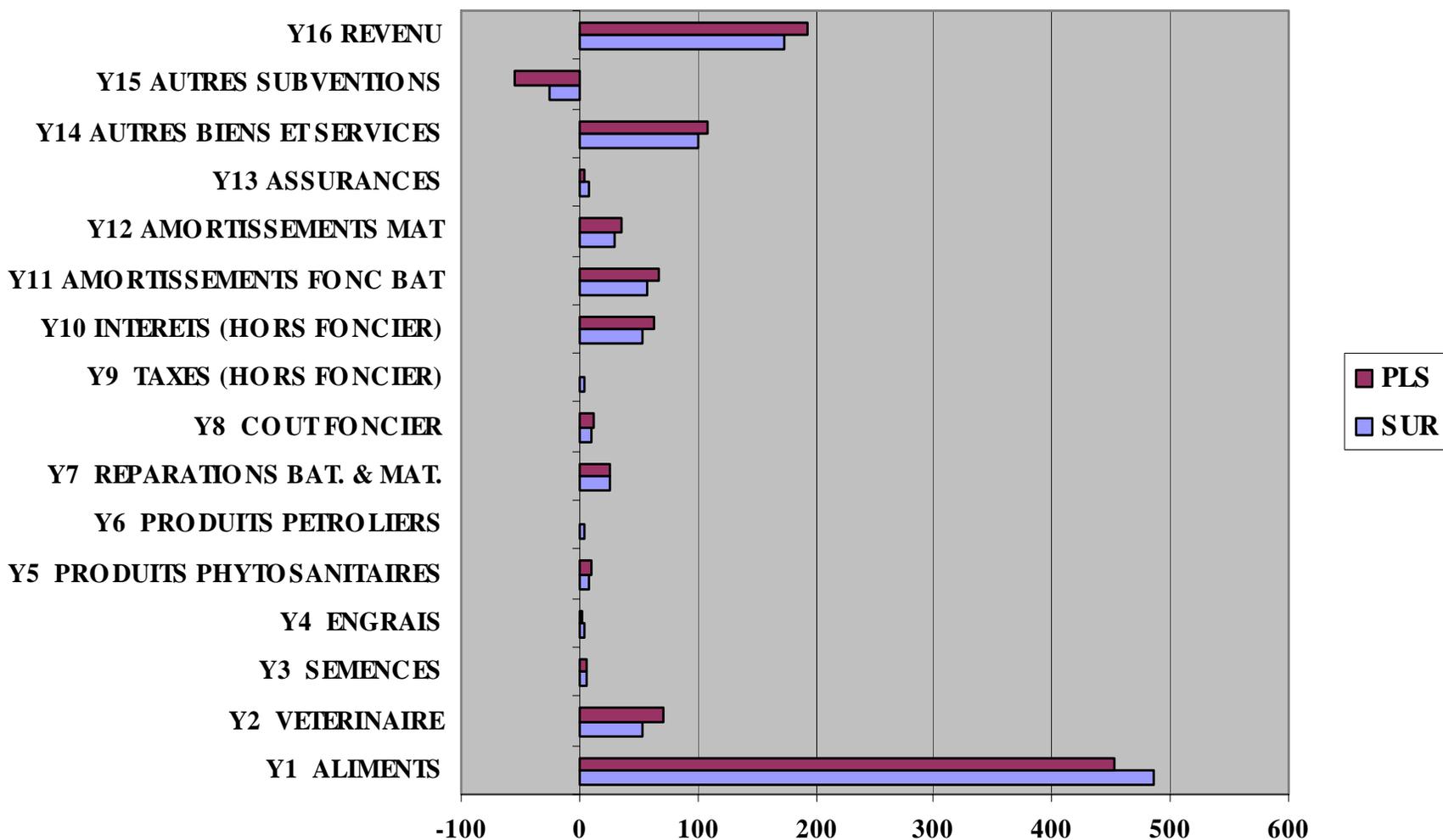
X10 tournesol





Comparaison des résultats : le porc

X20 porc





Comparaison des résultats : le lait

X22 lait

