

Evaluer les programmes d'insertion et d'emploi

Bruno Crépon, CREST J-PAL

Que veut dire évaluer?

- ❑ Souvent chercher à répondre à des questions sur la mise en œuvre du programme
 - Combien de bénéficiaires? / Quel coût? /Ce qui était prévu a-t-il été fait?
- ❑ Utile mais insuffisant :
 - Les programmes sont conçus par rapport à un objectif.
 - Evaluer c'est aussi examiner si les programmes atteignent leurs objectifs
- ➡ Evaluer l'**impact** du programme

Exemple : Micro-crédit

- Problème : Absence de services financiers pour les plus pauvres dans les PVD (et ailleurs).
 - Limite la capacité de développer une activité
 - Prive beaucoup qui auraient la capacité d'être entrepreneur d'entreprendre
- Programme : rendre l'accès aux prêts aux plus pauvres possible – Institution de micro-finance
- Questions :
 - L'accès aux services financiers permet-il de réduire la pauvreté
 - Quelles sont les conséquences sur la consommation, les activités, l'investissement...

Exemple : Opérateurs Privés de Placement et Cap Vers l'Entreprise

- ❑ Demandeurs d'emploi avec risque de chômage de longue durée à l'inscription
- ❑ Portefeuille d'un agent de l'ANPE : 120 DE.
- ❑ Pour le DE dans les programmes OPP et CVE 40 DE par portefeuille.
- ❑ Objectif : améliorer le retour à l'emploi dans un contexte plus efficace
- ❑ Questions :
 - La réduction de la taille des portefeuilles accroît-elle le retour à l'emploi
 - Vaut-il mieux un accompagnement public ou un accompagnement privé

Exemples : Les jeunes et la santé

- Problème :
 - Mauvais état de santé des jeunes des banlieues défavorisées / faible recours aux soins / absence de “culture” de la santé
- Programme : faire bénéficier les jeunes des ML
 - Consultations avec un médecin et un psychologue en ML
 - Information / assistance sur les droits
- Questions :
 - Les consultations améliorent-t-elles le recours aux soins / le rapport à la santé / impact sur l’insertion
 - Fournir des informations et une assistance sur les droits est-il substituable?

Pourquoi est-ce important d'évaluer l'impact d'un programme?

- ❑ On ne sait pas distinguer ce qui marche de ce qui ne marche pas
- ❑ Les praticiens peuvent avoir une idée tirée de leur expérience
 - Mais cette idée est dure à faire partager
 - Eux même peuvent se tromper
- ❑ Pas de consensus finalement sur
 - Quels sont les problèmes
 - Quelles sont leurs solutions
- ➡ Mauvaise allocation des fonds
 - Encore plus tendance à limiter les investissements et à saupoudrer sur un grand nombre de projets

Les Politiques Actives du Marché du Travail

- ❑ Un cas particulièrement emblématique
- ❑ Des années de pratiques
- ❑ Des années aussi de chômage important
- ❑ Pourtant pas encore aujourd'hui de consensus sur
 - Quelles sont les vrais besoins des demandeurs d'emploi : formation / accompagnement / expérience / insertion
 - Quels sont les programmes qui fonctionnent / comment les attribuer / comment les améliorer
- ❑ La persistance de débats sans fin : le sentiment d'être dans l'impasse
 - Parfois aussi du défaitisme : on a tout essayé !
 - Mais est-ce qu'on s'est seulement donné le moyen de faire le tri dans ce qu'on a essayé?

Les Politiques Actives du Marché du Travail

- ❑ Une approche très grossière. Souvent considérée comme un tout
 - ❑ Les politiques actives du marché du travail ont un contenu
 - ❑ Les formations par exemple ont des modules représentant l'idée que l'on se fait des besoins des demandeurs d'emploi
 - ❑ L'accompagnement a aussi des composantes correspondant à une idée que l'on se fait des besoins des demandeurs d'emploi
 - ❑ On n'a en fait aucune idée ni aucune compréhension réellement établie de ce que sont vraiment ces besoins et de la façon dont on les satisfait
-

Evaluer ne va pas de soi

- Difficulté fondamentale de l'évaluation
 - Comparer la situation d'un individu participant avec *ce qu'aurait été la situation de cet individu en l'absence du programme*
 - Difficile car soit il suit le programme.... soit il ne le suit pas
- On ne va pas s'en tirer simplement avec une enquête ou quelques interviews sur le ressenti
 - On peut toujours le faire mais c'est pour répondre à une autre question que celle de l'impact
- ➡ Nécessité de se donner les moyens d'évaluer
 - ➡ C'est à dire se donner les moyens de reconstituer cette information manquante : qu'aurait été la situation des individus s'ils n'avaient pas bénéficié du programme

Se donner les moyens d'évaluer mais pas n'importe lesquels : besoin d'évidence

- ❑ Souvent des a priori sur ce qui marche et ce qui ne marche pas
 - Des débats qui restent très idéologiques : Exemple Micro-crédit
 - ❑ Micro-crédit = merveille / tous les pauvres sont des entrepreneurs potentiels
 - ❑ Micro-crédit = spirale infernale du surendettement
 - Loin des réalités concrètes
 - ❑ Programmes sont redistributifs / les gens ne sont jamais d'accord sur l'utilisation des fonds
 - ❑ Des programmes dont les effets sont surement modestes
 - La première question à laquelle on veut répondre est "Est-ce que cela marche !"
- ➔ Besoin d'évidences, obtenues d'autant plus rigoureusement que les a priori sont forts

Comment évaluer?

- Idée générale : comparaison de bénéficiaires avec un groupe de contrôle non bénéficiaire

- On s'intéresse au changement apporté par la participation au programme
 - On ne s'intéresse pas aux caractéristiques de la population

- Notion centrale : toutes choses égales par ailleurs
 - Les deux populations bénéficiaires et non bénéficiaires doivent être statistiquement « identiques »
 - Le but de toutes les évaluations d'impact est de parvenir à faire de telles comparaisons

Des méthodes peu concluantes

- Comparer les bénéficiaires avec des non bénéficiaires
 - Comment les choisir?
 - Exemple : un programme qui concerne les femmes seules avec enfant. Quel est le groupe de contrôle : les femmes seules – les femmes avec enfants...

- Problème : population des bénéficiaires en général très particulière
 - Cas des jeunes voulant créer une entreprise
 - Comparaison entre bénéficiaires et non bénéficiaires mesure l'effet du programme + l'effet population

- Problème dit du **biais de sélection**
 - L'élimination du biais de sélection est la principale difficulté des évaluations d'impact

La méthode du matching

- ❑ Appariement de chaque bénéficiaire avec un « jumeau » n'en bénéficiant pas
- ❑ Jumeau: même âge, diplôme, historique, sexe, etc.,
- ❑ Impact = différence de situation moyenne entre les jumeaux.
- ❑ Une idée simple naturelle et attractive,...
- ❑ De nombreuses déclinaisons du simple au très sophistiqué
 - Régression à variable de contrôle
 - Matching simple
 - ❑ De nombreuses sophistications de la technologie du matching
 - Méthode d'approximation polynomiale
 - Propensity score matching
 - ❑ 1 Logit expliquant la décision de participer
 - ❑ 2 Appariement des individus sur le score et non sur l'ensemble des caractéristiques
 - De nombreuses sophistications aussi pour le calcul des écart-types
 - ❑ Méthodes du bootstrap par exemple

Des raffinements somme toute peu efficaces

- Une idée qui marche,... mais pas suffisamment bien
 - Rend les comparaisons entre bénéficiaires et non bénéficiaires plus pertinentes,
 - Mais il reste encore des différences dans la comparaison entre bénéficiaires et non bénéficiaires
 - **Le biais résiduel est en général de l'ordre de grandeur du programme**

- Il faut être modeste : les programmes que l'on développe ont des effets de faible ampleur
- Mais il faut être exigeant : on veut savoir ce qui marche et ce qui ne marche pas : on doit utiliser des méthodes parfaitement rigoureuses
 - Il ne s'agit pas de se faire plaisir en utilisant des méthodes indument raffinées

L'évaluation par expérimentations aléatoires

- ❑ Deux populations tirées au hasard, une bénéficiant du programme et l'autre non
- ❑ La population des non bénéficiaires est un groupe de contrôle idéal pour évaluer l'effet du programme
- ❑ Il n'y a plus d'effet population
- ❑ L'effet du programme est mesuré en toute transparence par la différence de situations de deux populations construites pour être comparables.
- ❑ Une méthode fiable transparente et précise lorsque les échantillons sont suffisamment importants

Est-ce que c'est possible?

- ❑ Peut paraître complexe et peu opérationnel
- ❑ En fait, plutôt un principe qu'il est possible d'insérer de façon souple, dans le processus de mise en œuvre d'un programme.

De très nombreuses façons d'organiser le tirage au sort

- Tirer des individus intéressés lorsqu'il y a plus de bénéficiaires que de places disponibles
- Tirer des individus éligibles et donner une incitation à entrer dans le programme (parfois les bénéficiaires ne sont pas si intéressés par le programme)
- Tirer des groupes d'individus, par exemple des missions locales

OPP et CVE : Nombre de places et tailles des populations éligibles

- ❑ 40.000 places pour OPP, 40.000 places pour CVE
- ❑ 250.000 bénéficiaires potentiels
- ❑ Outil de Constitution des Cohortes développé dans 270 Agences Locales de l'Emploi (pas la totalité des agences)
- ❑ A l'inscription en chômage un DE avec risque de chômage de longue durée est envoyé aléatoirement dans trois programmes possibles :
 - L'accompagnement OPP
 - L'accompagnement CVE
 - Le programme classique d'accompagnement de l'ANPE
- ❑ Constitution de trois populations statistiquement identiques

Garantie de Ressource

- ❑ Problème des contraintes financière des jeunes et de leur participation active à leur insertion :
 - Donner une garantie de ressource tant qu'il y a participation du jeune
- ❑ Identifier les missions locales intéressées par le programme :
 - Les organiser par paires
 - Tirer au sort celles qui seront traitement et celle qui seront contrôle
 - Tous les jeunes de la mission locale test bénéficient de la grantie
 - En pratique très simple à mettre en place
- ❑ Marche bien si il y a un grand nombre de Missions Locales

Où en sommes nous en France

- ❑ Longtemps à la traîne de la question de l'évaluation
- ❑ Les choses ont bien changé depuis
- ❑ La France est maintenant probablement à la pointe en matière d'évaluation et en particulier en matière d'évaluation randomisée
- ❑ Plusieurs facteurs ont contribué
 - Dynamisme de Pôle Emploi et de la Dares
 - Révolution du fonds d'expérimentation de la jeunesse
 - L'implantation et le rôle fédérateur du Poverty Action Lab

Dynamisme du SPE

- Evaluation randomisée des programmes d'accompagnement renforcés
 - Pôle Emploi : chômeurs à risque de chômage de longue durée
 - Dares : désir de s'approprier résolument la technologie de ces méthodes : jeunes diplômés au chômage depuis plus de un an
- Politique d'anonymisation des cv confiée par le gouvernement à Pole Emploi
 - Qui décide de nous confier son évaluation par évaluation randomisée

La révolution du Fonds d'expérimentation de la jeunesse

- ❑ Actuellement un grand nombre de programmes évalués dans le cadre du Fond d'Expérimentation pour la Jeunesse
 - Idée générale
 - ❑ Ne nous privons pas d'avoir de bonnes idées
 - ❑ Mais soumettons nous à la discipline de les évaluer
- ❑ Montant important (150 millions d'euros) consacré à la mise en oeuvre de nouveaux projets et à leur évaluation
- ❑ Moyen puissant d'accumuler de l'expérience sur ce qui fonctionne et ne fonctionne pas et sur où sont les vrais problèmes
- ❑ Idée importante de partir du terrain

La révolution du Fonds d'expérimentation de la jeunesse

- Santé...
 - Problématique du non recours
- Logement
 - Problème de la caution
- Education/orientation
 - Rupture en apprentissage
- Permis de conduire
 - Financer le permis de conduire de 10.000 jeunes
- Création d'entreprise
 - Développer la conduite de projet

Le Poverty Action Lab

- ❑ Créer par des chercheurs du MIT (dont Esther Duflo) pour promouvoir les évaluations randomisées
- ❑ Forte implication dans le domaine des pays en développement
- ❑ Maintenant un laboratoire fédérant plus de 50 chercheurs dans le monde
- ❑ Nombreux thèmes dont l'emploi et les politiques à destination des jeunes

Le Poverty Action Lab

- ❑ Antenne à Paris
- ❑ Contribuer à créer une offre structurée permettant de répondre à la demande des pouvoirs publics
- ❑ Accumulation d'une expérience et d'un savoir faire dans la conduite d'évaluations randomisées
- ❑ Un savoir qui se diffuse et fait maintenant référence

Pour finir : un exemple

Programme d'accompagnement renforcé des demandeurs d'emploi OPP/CVE

Design : tirage au sort bénéficiaires potentiels / contrôle potentiel

Bénéficiaires potentiels peuvent refuser le programme

Illustration OPP

- ❑ On considère l'évaluation OPP
- ❑ On montre les résultats obtenus avec l'évaluation randomisée
- ❑ Dans le groupe OPP il y a les acceptants et les refusants
 - On fait aussi l'évaluation en ignorant le groupe de contrôle expérimental
 - On utilise seulement les refusants comme contrôle pour les acceptants
- ❑ Question : est-il nécessaire de mettre en œuvre les méthodes expérimentales pour évaluer
- ❑ Ne peut-on pas se contenter de méthodes non-expérimentales?

OPP : Evaluation expérimentale

TABLE 1 – Principaux résultats de l'expérience aléatoire pour le programme OPP

	Effet	I.C.
Taux de sortie à 3 mois (%)	1,6	[-1,1;4,3]
Taux de sortie à 6 mois (%)	4,2*	[0,8;7,6]
Taux de sortie à 12 mois (%)	5,6**	[1,6;9,6]
Jours d'inscription au chômage	-0,8	[-8,9;7,3]
Effectif	63 390	

Significance levels : † : 10% * : 5% ** : 1%

Méthode non expérimentales

- ❑ On se concentre sur la variable de durée en chômage
- ❑ On considère les méthodes type matching
 - Regression a variable de contrôle / propensity score matching
- ❑ Appliquée aux groupes entrant OPP et refusant OPP
- ❑ Différents jeux de variables de contrôle
- ❑ Question : est-il nécessaire de mettre en œuvre les méthodes expérimentales pour évaluer
- ❑ Ne peut-on pas se contenter de méthodes non-expérimentales?
- ❑ Jusqu'à 74 variables de contrôle

Evaluation par régression

Table 9: Effet de la variable *AcceptationOPP*

Contrôles	Coefficient	(Std. Err.)	R ²
<i>Aucun</i>	27.27**	(1.32)	0.032
<i>Caractéristiques du chômeur</i>	25.67**	(1.29)	0.081
<i>Historique de chômage du chômeur</i>	27.31**	(1.30)	0.066
<i>Motivation du chômeur</i>	29.18**	(1.32)	0.050
<i>Caractéristiques et Historique du chômeur</i>	26.23**	(1.29)	0.097
<i>Caractéristiques et Motivation du chômeur</i>	26.97**	(1.30)	0.092
<i>Historique et Motivation du chômeur</i>	29.24**	(1.30)	0.080
<i>Tous</i>	27.52**	(1.29)	0.107

Significance levels : † : 10% * : 5% ** : 1%

Les écart-types sont calculés à l'aide la méthode de White

Toutes les régressions incluent des effets fixes par ALE et contrôlent par le trimestre d'inscription à OCC entre 2006 et 2007.



Des méthodes plus sophistiquées

TABLE 5 – Résultat du matching sur le score de propension

Méthode utilisée	ATT	S.E
Commande Stata : Psmatch2		
Plus proche voisin	27.30	2.20
5 plus proches voisins	27.60	1,87
Commande Stata : nnmatch		
Plus proche voisin sans biais	27.30	2.20
Plus proche voisin, écart-types robustes sur 1 match	27.30	2.27
Plus proche voisin, écart-types robustes sur 5 match	27.30	2.32
Plus proche voisin sur le pscore et Exact sur 3 variables	28.75	1.94

Le support commun est défini par une méthode de trimming à 10 %



Conclusion

- ❑ Evaluer est une **nécessité** pour améliorer les politiques publiques
- ❑ Evaluer de façon rigoureuse, fiable et précise est **possible**
- ❑ Il est nécessaire pour parvenir à des évaluations satisfaisantes de **s'en donner les moyens**
 - Les méthodes non randomisées conduisent le plus souvent à des évaluations très éloignées de la vérité
 - Les méthodes d'évaluations randomisées peuvent être mises en œuvre facilement dans un grand nombre de situations

Conclusion

- Nombreuses évaluations randomisées conduites en France
 - Une offre existante via J-Pal
 - La création du fonds d'expérimentation de la jeunesse a été une initiative très innovante
- Qu'en est-il des politiques du marché du travail?
 - Est-ce que ce ne serait pas une bonne idée de créer un fonds d'expérimentation des politiques de l'emploi?



Principes généraux d'évaluation des effets d'un dispositif

Bruno Van der Linden

Fonds de la Recherche Scientifique

Institut de Recherches Economiques et Sociales
U.C.L.

<http://perso.uclouvain.be/bruno.vanderlinden/>





INTRODUCTION



Pourquoi évaluer?

1. Les ressources *collectives* sont rares
2. par comparaison avec l'étendue des besoins: pauvreté, chômage, inégalités, dégradation du climat, ...
3. compte tenu du coût des « dispositifs », « politiques », « programmes », « projets », « actions »,...

Conviction: On connaît étonnamment *peu* sur « ce qui marche vraiment » parce qu'on évalue mal.



Organisation des thèmes

1. Evaluation des besoins et programmation (avant le démarrage de l'action)
2. Evaluation du processus
3. Evaluation d'impact

Evaluations du processus et d'impact
pas seulement quand l'action a pris fin
aussi en cours d'action (on going, in itinere)





APERÇU RAPIDE DES DEUX PREMIERS TYPES D'ÉVALUATIONS



L'évaluation des besoins et la programmation (en résumé)

1. Conception et objectifs du programme
 - > Quels sont les besoins (diagnostic!)?
 - > Quels sont les objectifs ?
2. Elle choisit une action compte tenu du primo
 - > Les actions alternatives doivent être connues
 - > Bonne idée de la nature des bénéfices attendus des diverses actions
3. Elle définit une population cible

L'évaluation de processus (en résumé)

1. L'action mise en place correspond-elle aux objectifs?
2. Livre-t-on effectivement les « services » annoncés au public prévu?
3. Peut-on améliorer l'efficacité ?
4. A quel degré l'action est-elle utilisée? Degré de satisfaction des utilisateurs?
5. Contrôle comptable et financier.



EVALUATION D'IMPACT



1. Objectifs de l'évaluation d'impact

1. Question fondamentale: L'action a-t-elle un impact (effet ou *résultat*) sur un indicateur (étroitement lié aux objectifs de celle-ci)?
2. Questions auxiliaires:
 - > Sur quelle(s) population(s) y a-t-il un impact? Idem pour l'absence d'impact?
 - > Peut-on préciser les caractéristiques de l'impact?
 - » Exemple: Impact avéré sur l'emploi
 - » Au-delà: durée de l'emploi? Salaire ? ...
 - > ...

2. Notions centrales

Mesurer l'impact nécessite une analyse de **causalité**

Exemple :

- > le nombre de stagiaires d'une formation embauchés = une **réalisation** qu'on peut facilement *observer*
→ « *taux de placement* »
≠ une analyse de causalité
≠ une évaluation !!!
- > le nombre de stagiaires embauchés **grâce à** la formation = un **résultat** qu'on ne peut pas observer car l'embauche pourrait aussi avoir eu lieu en l'absence de formation
→ « *question centrale de l'évaluation* » ...

Notions centrales

La question centrale de l'évaluation:

Quel est le résultat moyen d'une action sur les bénéficiaires de l'action ?

- > **résultat** = *variation* de l'indicateur *grâce* à l'action
 - » Indicateur (*insertion en emploi, salaire,...*)
 - » Variation grâce à l'action = différence entre
 - Le niveau de l'indicateur après l'action;
 - Le niveau de l'indicateur *au même moment* **en l'absence de l'action**
 - » *Exemple* :
variation du taux d'embauche des stagiaires *grâce* à la formation

- > **sur les bénéficiaires** (= micro <-> macro)



Notions centrales

Le problème fondamental de l'évaluation

L'indicateur *en l'absence de l'action* ne peut pas être observé, car on ne peut *pas simultanément participer à une action et ne pas y participer !*



Notions centrales

Solution au problème fondamental

- > Trouver un groupe de non-participants pour *approcher* le niveau de l'indicateur *en l'absence* de l'action
- > **Problème : le biais de sélection**
- > **Un tel biais se présente lorsque simultanément**
 1. Caractéristiques individuelles participants \neq non-participants
 2. Certaines de ces caractéristiques ne sont pas observées
 3. Ces caractéristiques sont *corrélées* avec l'indicateur
- la différence entre l'indicateur des participants et des non-participants reflète leurs caractéristiques différentes
- **biaise** la comparaison → **biaise** la mesure de l'impact
- > **Recherche de solution au biais de sélection = au centre de toute la suite de l'exposé**

3. Comment évaluer l'impact?

Différentes méthodes d'évaluation :

= différentes définitions de la « *situation de référence* » définie ainsi:

« *ce qui serait advenu aux mêmes bénéficiaires en l'absence de l'action* »

→ requiert des hypothèses dont il faut vérifier la crédibilité au cas par cas!

Survolons les méthodes principales...



A. Reconstitution subjective

= demander directement aux bénéficiaires ou à ceux qui les ont recrutés « ce qui serait advenu en l'absence de l'action ».

- > *Avantages* :
pas de problème de biais de sélection
- > *Inconvénients* :
subjectivité, risque d'erreurs, voire manipulation

Bilan: Méthode à déconseiller.

B. La méthode avant - après

Situation de référence :

mesurée par la réalisation de l'indicateur de résultat parmi les *bénéficiaires avant* participation

> *Avantages :*

pas de biais de sélection car l'on compare les mêmes individus (les bénéficiaires avant et après l'action)

> *Inconvénients :*

» Ne permet pas de distinguer clairement l'effet de l'action de celui d'autres phénomènes variant au cours du temps (*modification de la conjoncture économique par exemple*)

» Risque de biais si effets d'anticipation présents

Bilan: Très rarement une approche valide.

C. Groupe témoin ou « de contrôle »

Situation de référence :

mesurée par la réalisation de l'indicateur de résultat parmi un groupe de **non bénéficiaires**
(= « groupe de contrôle » ou « groupe témoin »)

> *Avantages :*

pas de biais causé par la présence d'autres phénomènes variant dans le temps, car la comparaison se réalise au même moment = *après* l'action

> *Inconvénients :*

toujours un risque de biais de sélection

Bilan: à n'utiliser que dans certains cas (voir « comment constituer un groupe témoin? »)

D. Méthode de « différence de différences »

= Combinaison de la méthode avant-après et groupe témoin

Requiert un groupe témoin observé en deux points du temps : avant l'action et après celle-ci

Situation de référence :

mesurée par la réalisation de l'indicateur de résultat parmi un groupe de *non bénéficiaires* (= groupe de contrôle) *avant* et *après* l'action

Méthode à suivre:

1. Collecter des données initiales pour chaque groupe avant l'introduction du programme.
2. Collecter les données de suivi pour chaque groupe après que le programme a eu lieu.
3. Calculer la différence avant-après pour chaque groupe
4. Soustraire la différence observée dans le groupe témoin à celle observée dans le groupe traité

Groupe	Avant	Après	Différence Avant-Après
1. Bénéficiaires	2.0	3.0	1.0
2. Non bénéfic.	2.2	2.3	0.1
3. Impact			$1.0 - 0.1 = 0.9$

Méthode de « différence de différences »

1. Reconnaît que le groupe témoin peut-être différent au départ
2. Regarde la trajectoire du groupe témoin dans le temps pour prédire la trajectoire du groupe de bénéficiaires
3. Repose sur une hypothèse forte:
 - › Elle suppose que la trajectoire du groupe traité aurait été parallèle à la trajectoire du groupe témoin en l'absence du programme

Bilan: Fonctionne bien si cette hypothèse est crédible. Sinon, il reste des biais.

4. Comment constituer un groupe témoin?

1. L'approche usuelle = constituer ce groupe après l'action

- > Danger maximum de biais de sélection
- > Danger d'autant plus faible
 - » qu'on a une *masse* d'information sur les bénéficiaires et non bénéficiaires
 - » Un très grand choix possible d'individus non bénéficiaires

si bien qu'on peut apparier un bénéficiaire avec un (des) « quasi-jumeau(x) »

Note: choix de ces quasi-jumeaux requiert des méthodes statistiques sophistiquées, dites de « matching ».

Comment constituer un groupe témoin?

2. « Les expériences sociales contrôlées » : constituer le groupe **avant** l'action.

Plus précisément:

- » Identifier un ensemble de personnes qui satisfont toutes aux critères de sélection établis pour l'action.
- » Aléatoirement choisir la moitié des personnes qui participera à l'action (**groupe de « traitement »**);
- » L'autre moitié (**groupe de « contrôle »**) participera au programme plus tard (ou jamais) et, entre temps, servira de point de comparaison;
- » Après cette sélection, l'action prend place et on mesure les indicateurs de résultat sur les deux groupes (temps!)
- » La comparaison des impacts est simple.



Comment constituer un groupe témoin?

« Les expériences sociales contrôlées »

L'assignation aléatoire permet d'avoir un groupe témoin vraiment similaire au groupe de traitement

C'est la méthode d'évaluation la plus fiable (moyennant certaines précautions)

Note: l'assignation aléatoire peut prendre une multitude de formes (permettant d'échapper à la critique "cette méthode est injuste").



Comment constituer un groupe témoin?

3. Les « expériences naturelles »

Principe:

Utiliser la *réglementation de la politique* ou un *changement de politique* comme une « expérience »

Un exemple:

La réglementation peut créer une démarcation abrupte entre le public éligible et non éligible (seuil d'âge, seuil d'ancienneté, nombre d'enfants à charge, etc.) de manière permanente ou suite à un changement de la législation

Note: Requiert une analyse fine des conditions créées par la réglementation avant de l'assimiler à une expérience.



CONCLUSION



La question posée

1. Cet exposé s'est limité à quelques principes de base de l'évaluation « micro » entendue comme répondant à la question: « Quel est le résultat moyen d'une action sur les bénéficiaires de l'action ? »
2. D'autres questions pertinentes existent:
 - > Micro: il n'y a pas que l'effet moyen qui compte
 - > Macro: il n'y a pas que les bénéficiaires qui comptent. Ceux-ci peuvent dans certains cas s'en « sortir mieux au détriment de non-bénéficiaires ».

La mise en œuvre des méthodes

1. Pas *une* méthode qui s'indique toujours et partout
2. Message clé: évaluation ne se résume pas à une comparaison de phénomènes observés; pas d'évaluation sans hypothèses justifiées sur la manière de mesurer ce qui serait advenu en l'absence de la politique
3. Evaluer = un métier qui s'apprend à partir de connaissances de base (prérequis).

Les effets des allègements de cotisation sociale sur l'emploi et les salaires en France : une évaluation de la réforme Fillon de 2003

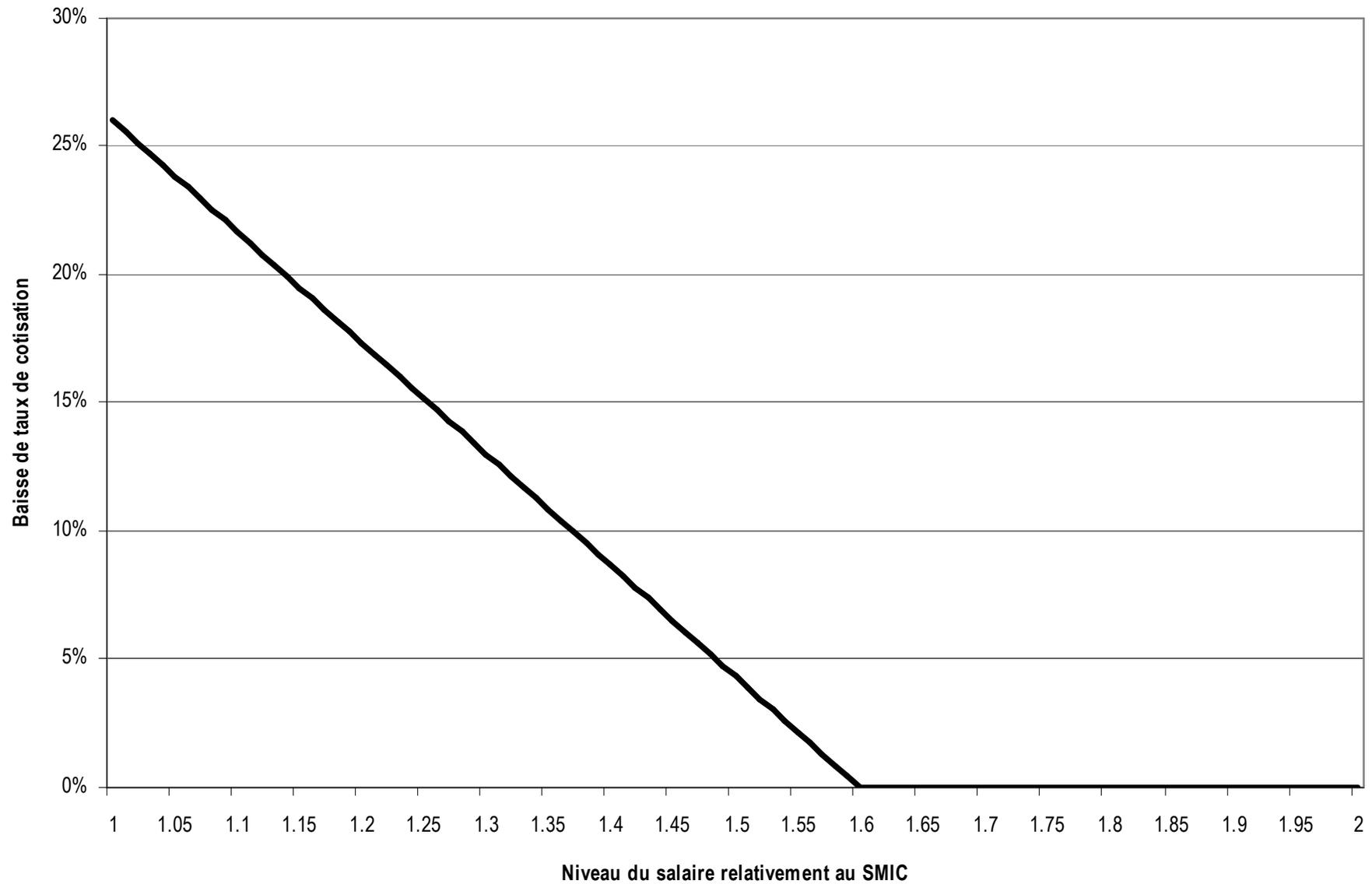
Matthieu BUNEL (*UTBM, CEE et TEPP*)

Fabrice GILLES (*EQUIPPE et CEE*)

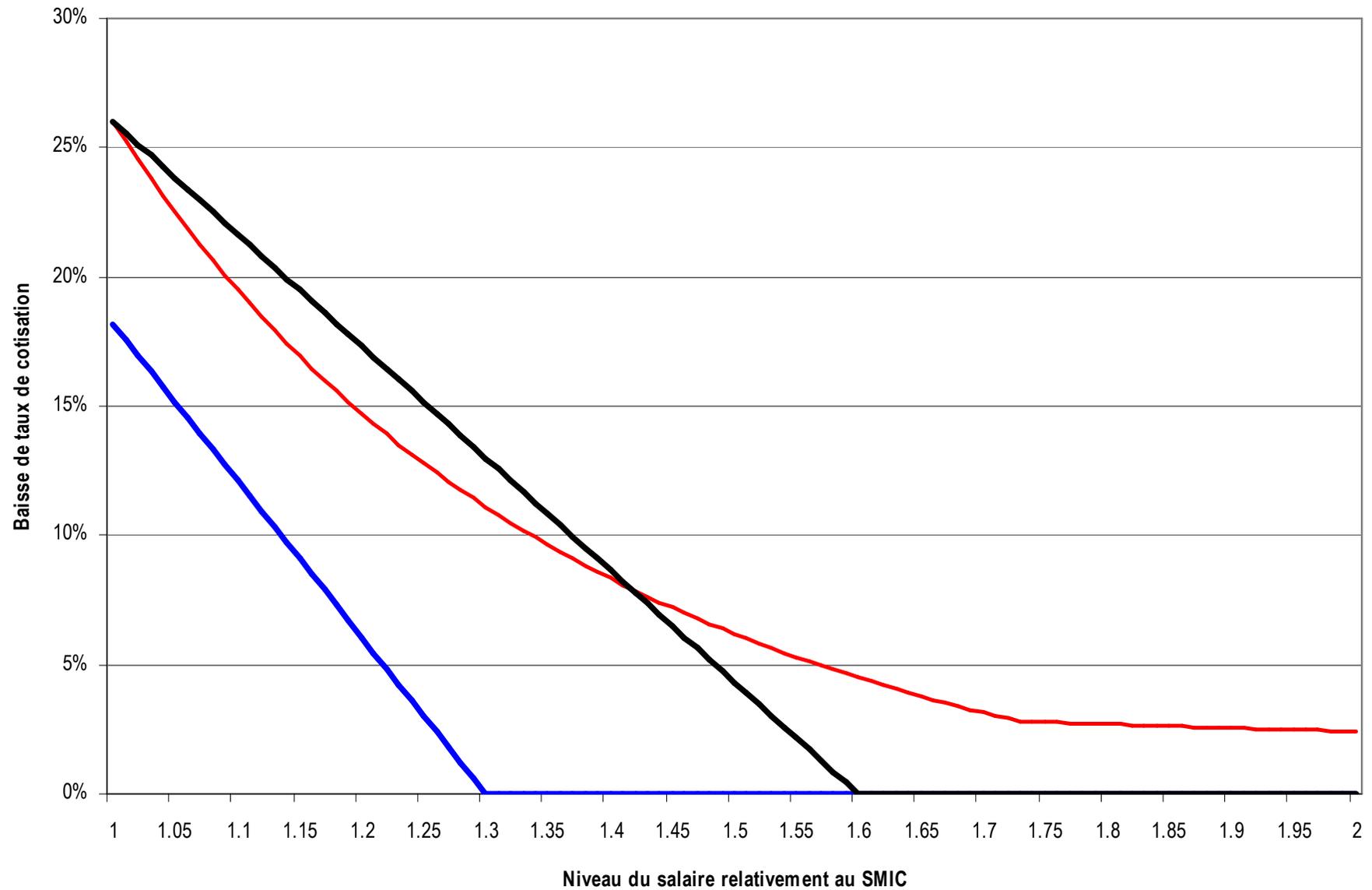
Yannick L'HORTY (*EPEE, CEE et TEPP*)



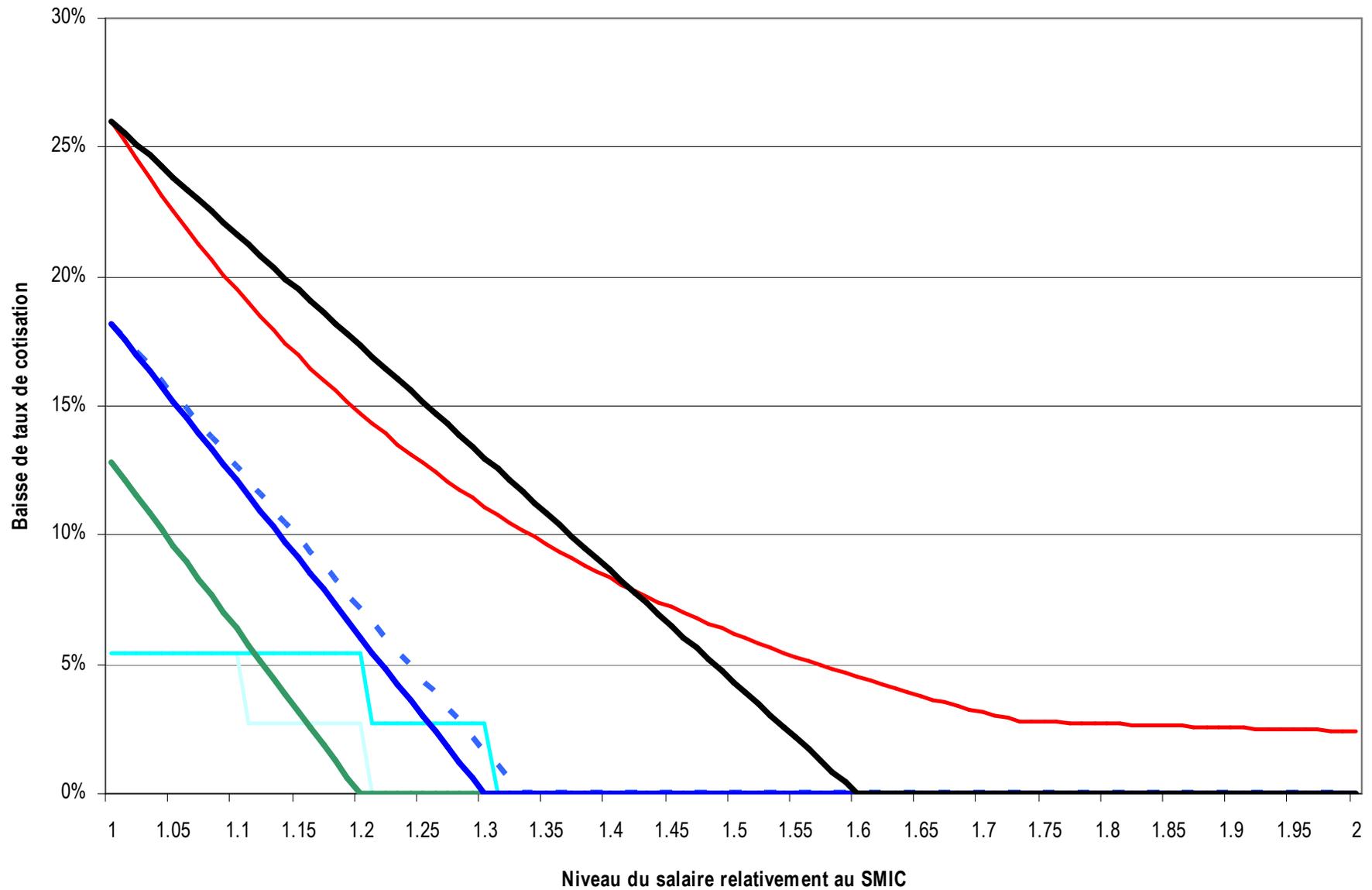
Le dispositif français d'allègements de cotisations sociales en 2010



La « réforme Fillon » de 2003-2005



3 générations de dispositifs



Enjeux de l'évaluation

- ❑ La première des politiques de l'emploi : 24,4 milliards d'Euros (2007) ;
- ❑ Une expérience française très variée
 - ❑ Depuis juillet 1993, toutes les formes de barèmes possibles ont été mises en œuvre, en métropole ou dans les DOM.
 - ❑ Au prix d'une grande instabilité (durée de vie des dispositifs : moins de 18 mois).
 - ❑ Les 3 records du Monde français: ...
- ❑ Qui s'inscrit dans une stratégie globale
 - ❑ Combiner un salaire minimum élevé avec une croissance riche en emploi
 - ❑ Activer le financement de la protection sociale au profit de l'emploi, sans pour autant faire le choix d'un barème progressif
 - ❑ « Agir à deux mains » sur le marché du travail : Subventionner les employeurs pour créer des emplois et les salariés pour qu'ils les occupent

Difficultés de l'évaluation

- ❑ Les réformes sont toujours un élément d'un mix plus complet
 - ❑ Avec la baisse de la durée du travail, en 1993, en 1998-2002
 - ❑ Avec des hausses du SMIC, en 1995, 1998-2002 et en 2003.

=> Un problème d'identification, voire même de dissociation

- ❑ Des mesures générales, qui touchent toutes les catégories d'entreprise

=> Pas de contrefactuel naturel ; pas de possibilité d'expérimentation

- ❑ Concentrées sur les bas salaires

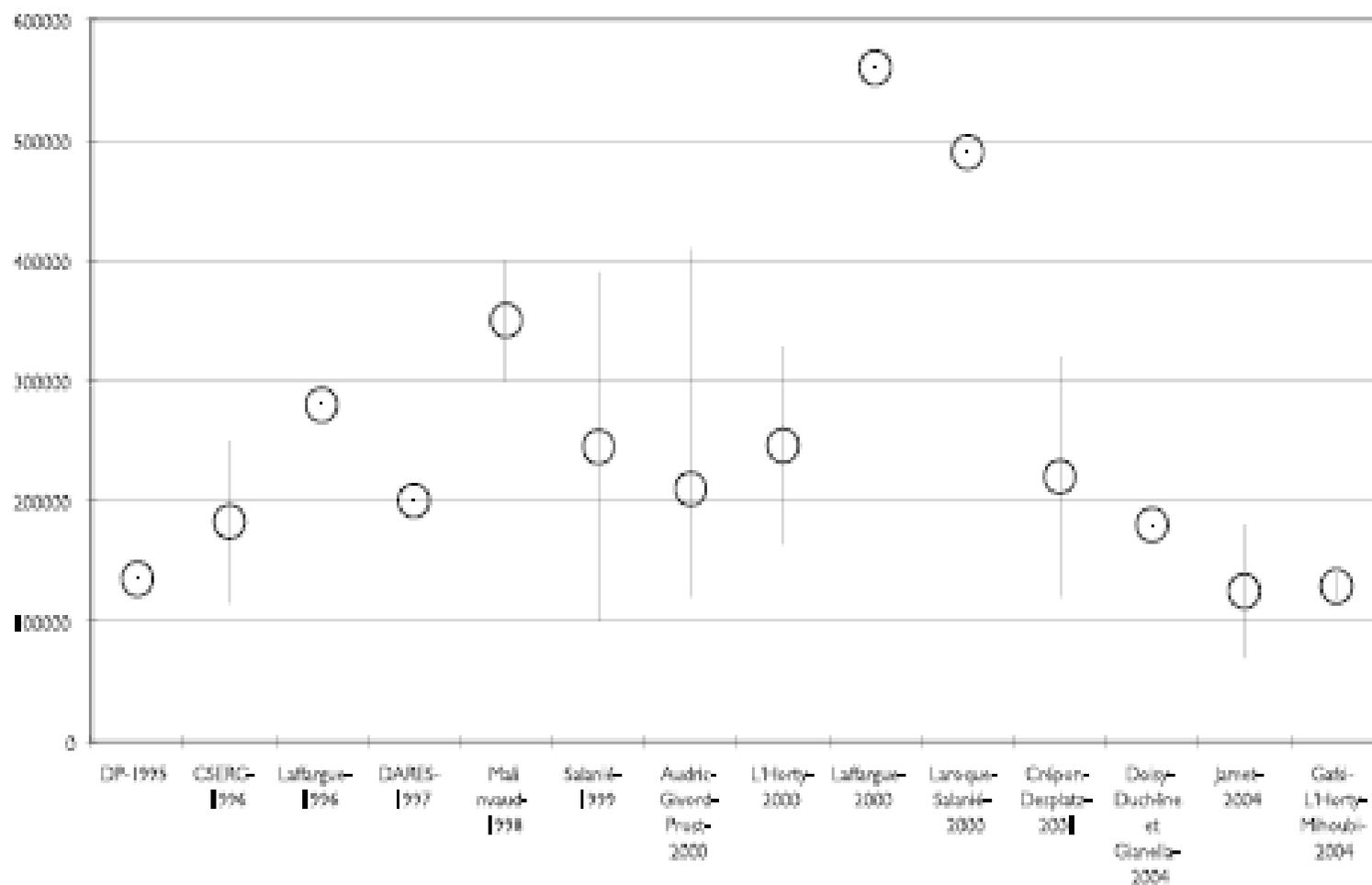
=> De l'hétérogénéité entre entreprises et entre les salariés

- ❑ Instables dans le temps

⇒ Pas de recul temporel dans les données entre deux chocs

⇒ Des différences de diagnostics entre les évaluations *ex ante* et *ex post*

NOMBRE D'EMPLOIS NON QUALIFIÉS CRÉÉS OU SAUVEGARDÉS SELON LES ÉVALUATIONS



Lecture : Les résultats des évaluations ont été ramenés à un budget de 5 milliards d'euros. Les traits continus correspondent aux intervalles de confiance des études, lorsqu'ils sont précisés.

Difficultés spécifiques pour la réforme de 2003-2005

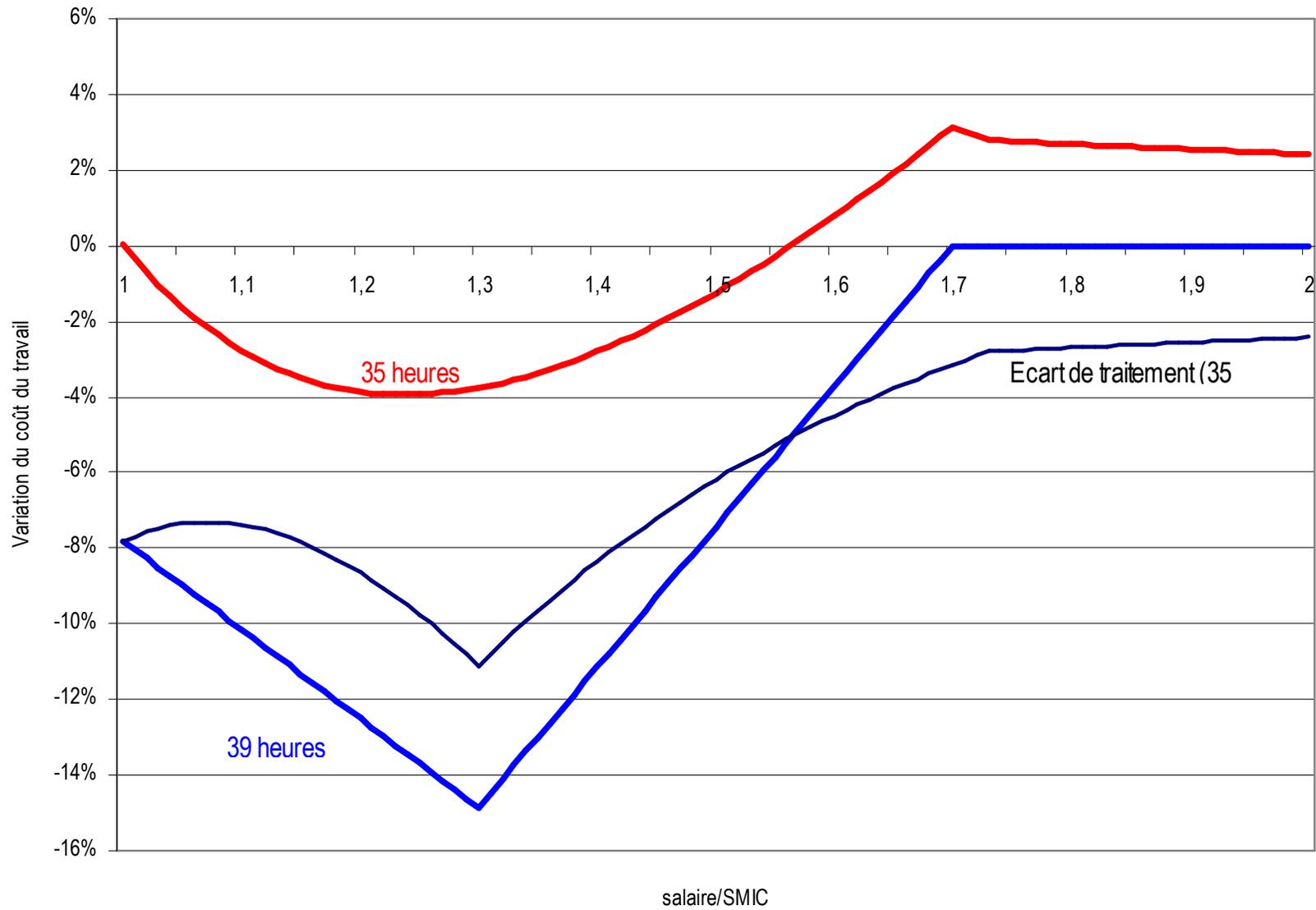
- Tous les problèmes précédents, sauf l'instabilité, depuis 2005 (mais données indisponibles).
- Une réforme progressive : pour les barèmes d'ACBS et pour les rémunérations minimales

=> Trois dispositifs intermédiaires entre 2003 et 2005

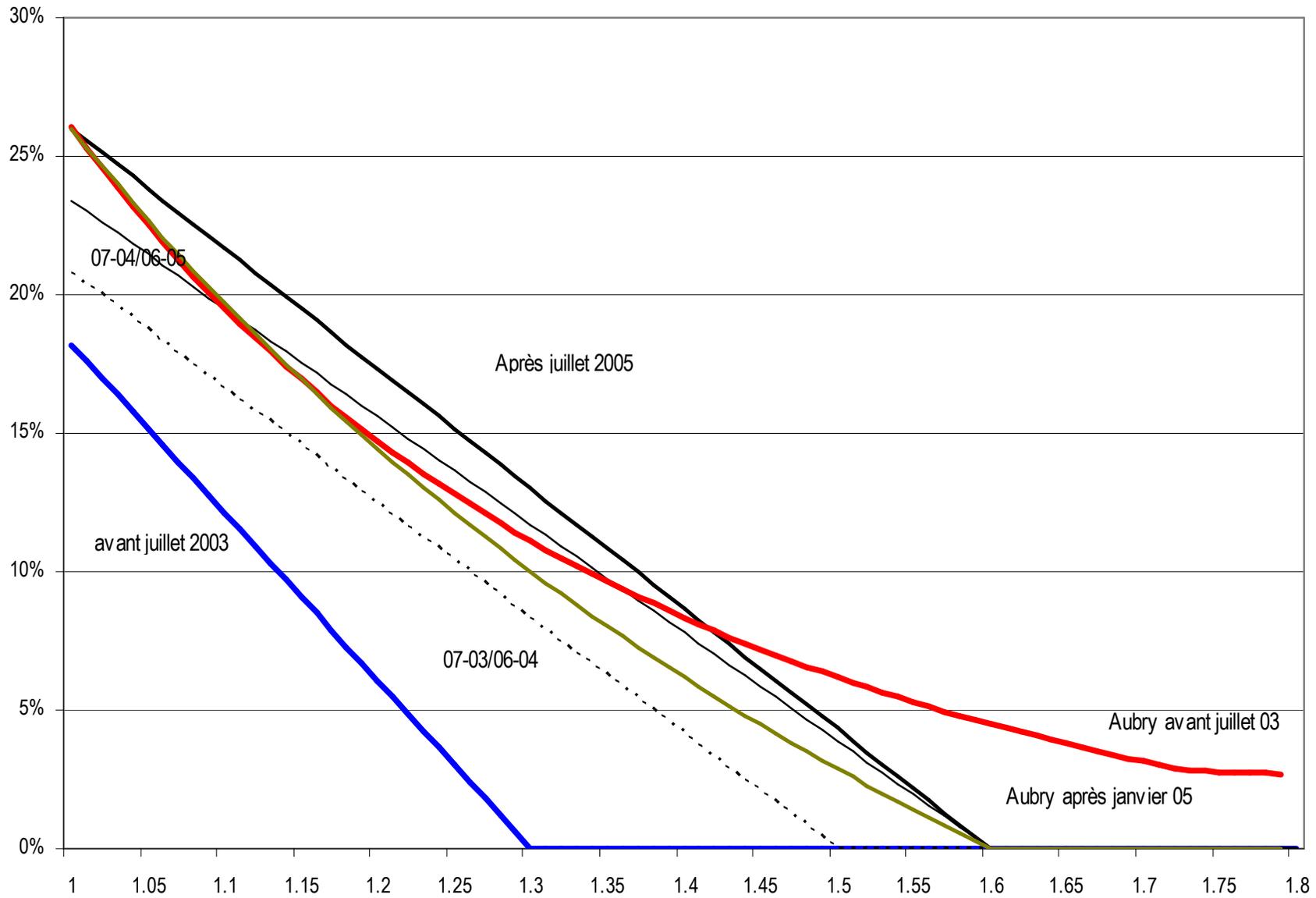
- Une réforme qui « harmonise »

Mais faire simple, c'est souvent compliqué

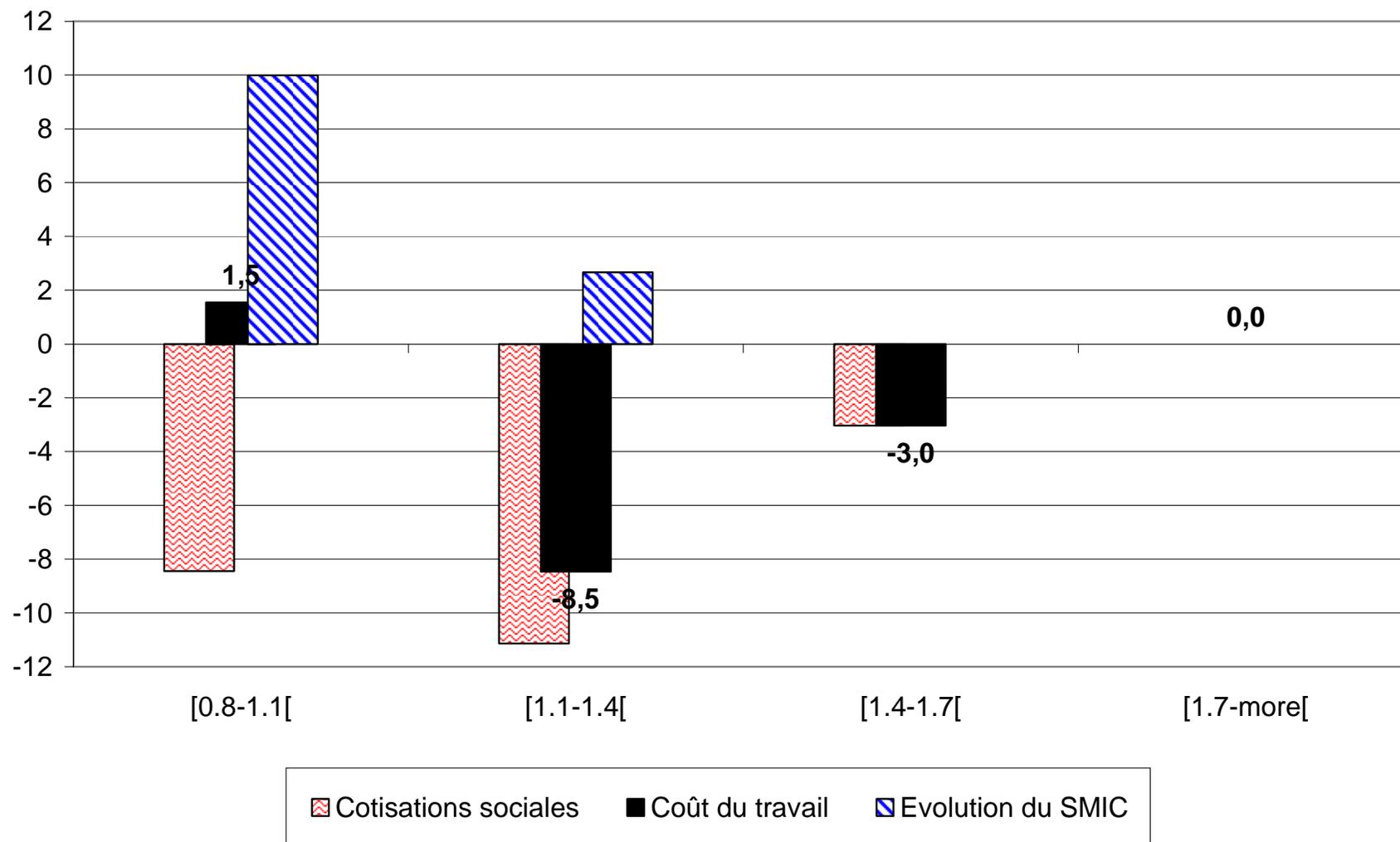
Anatomie du choc de cotisations



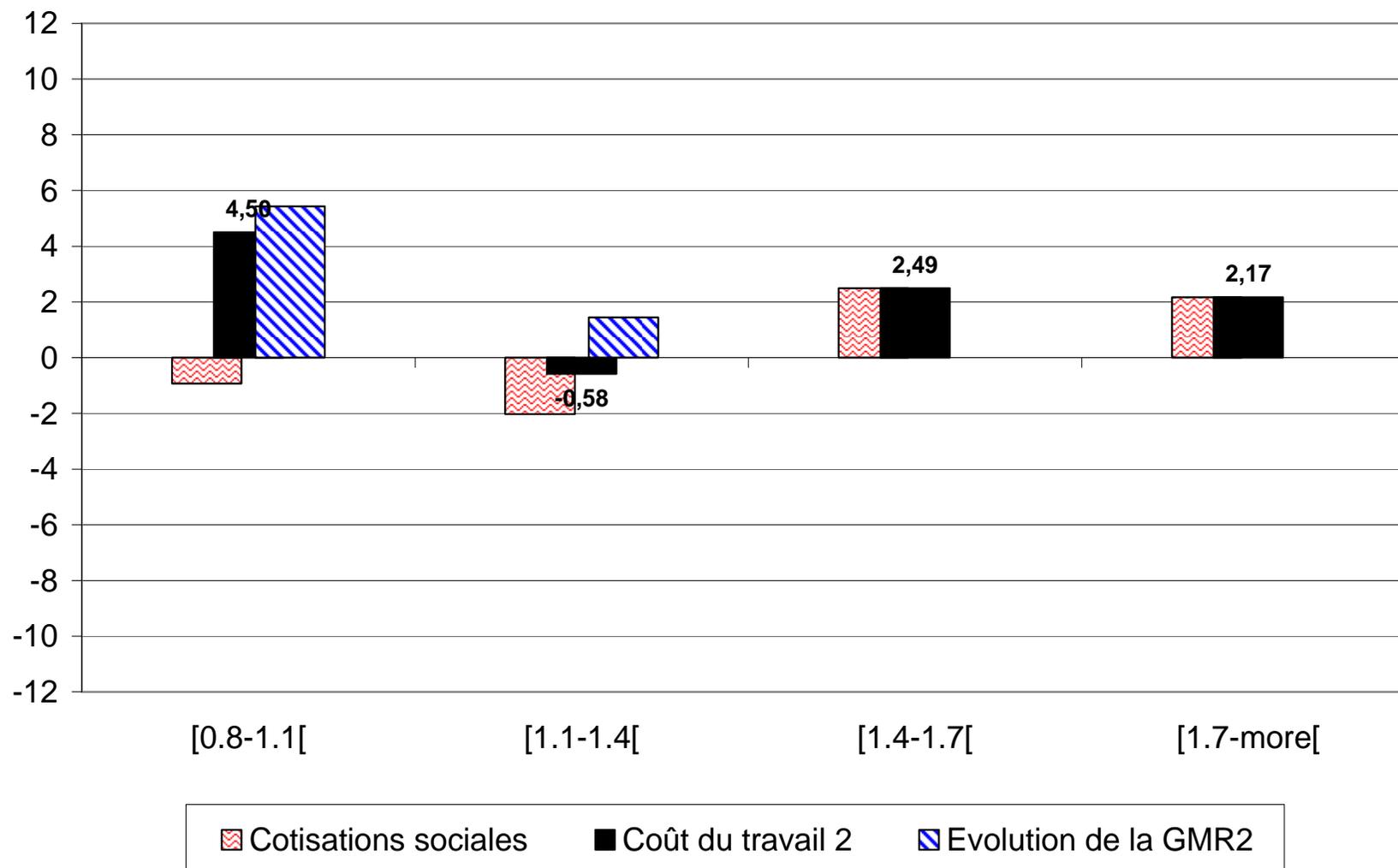
Les dispositifs intermédiaires



Variation du coût du travail : entreprises à 39 heures (Effets du passage du dispositif Juppé à Fillon)



Variation du coût du travail : entreprises à 35 heures (Effets du passage de Aubry 2 à Fillon)

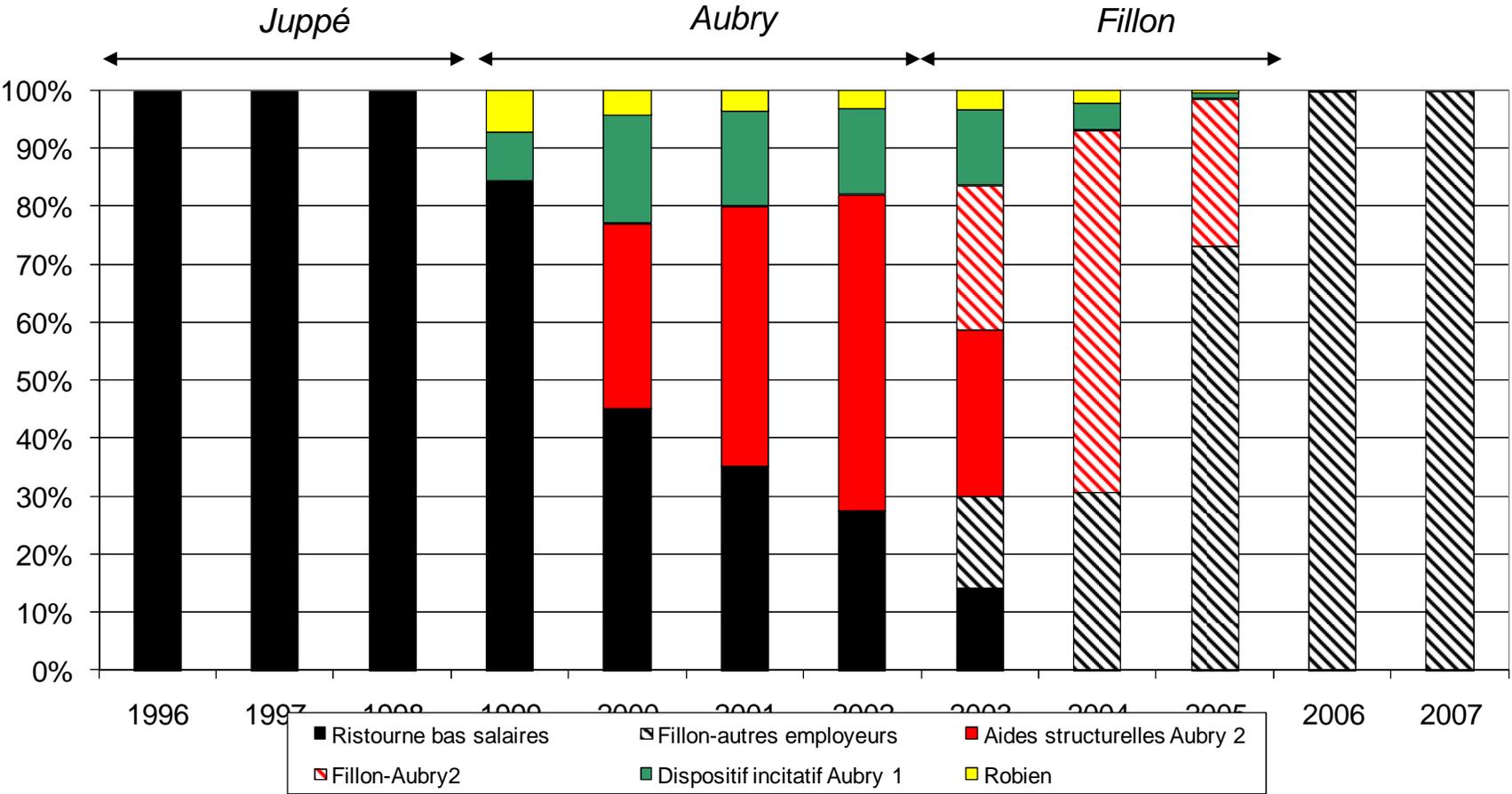


Les données

- ❑ Appariement de sources administratives exhaustives au niveau établissement
 - ❑ DADS, y compris fichiers poste (distribution de salaire intra-établissement)
 - ❑ SUSE, variables de contrôle niveau entreprise.
 - ❑ ACOSS, effectivité et montant des ACBS, niveau établissement.

- ❑ Champ de l'étude
 - ❑ Entreprises de plus de **5 salariés** (effectifs bruts renseignés au 31 décembre) issues du secteur privé non agricole et bénéficiant des allègements « Fillon après Aubry 2 » et « Fillon autres employeurs » de manière continue sur la période 2003 à 2004.
 - ❑ **Échantillon cylindré** sur la période 2002 à 2005 (bases Arome, Orme et Sequoia de l'Acoss). Les entreprises bénéficiant simultanément des deux types d'aides, les holdings, les entreprises de services domestiques, d'intérim ainsi que les entreprises publiques ont été supprimées.
 - ❑ L'échantillon final contient près de **75 000 entreprises** (95 000 établissements) employant 4 millions de salariés.

Distribution des établissements selon les dispositifs d'exonération



Stratégie d'estimation

$$\Delta \ln Y_i = X_i \beta + T_i \gamma + u_i$$

- ❑ Deux sous-échantillons
 - ❑ Entreprises à 35 heures : « Fillon après Aubry 2 »
 - ❑ Entreprises à 39 heures : « Fillon autres employeurs »

- ❑ Variable de traitement
 - ❑ Création d'une variable « Taux d'exonération apparent » au niveau états
 - ❑ Calcul de ce TEA à structure des salaires inchangée et en corrigeant des hausses du Smic et de la GMR

- ❑ Variable d'intérêt
 - ❑ Croissance de l'emploi (effectif, ETP),
 - ❑ Croissance des salaires et du coût du travail

- ❑ Méthodes d'estimation
 - ❑ MCO
 - ❑ Régressions quantiles
 - ❑ Scores de propension

Taux d'exonération apparents

Observé en 2005

$$\frac{Cotexo^{05}}{Sal^{05} + PP^{05}}$$

Calculé en 2005
sans correction de distribution des salaires

$$\frac{Cotexo(\widetilde{b}^{05}, dsal^{02})}{Sal^{02} + PP(\widetilde{b}^{05}, dsal^{02})}$$

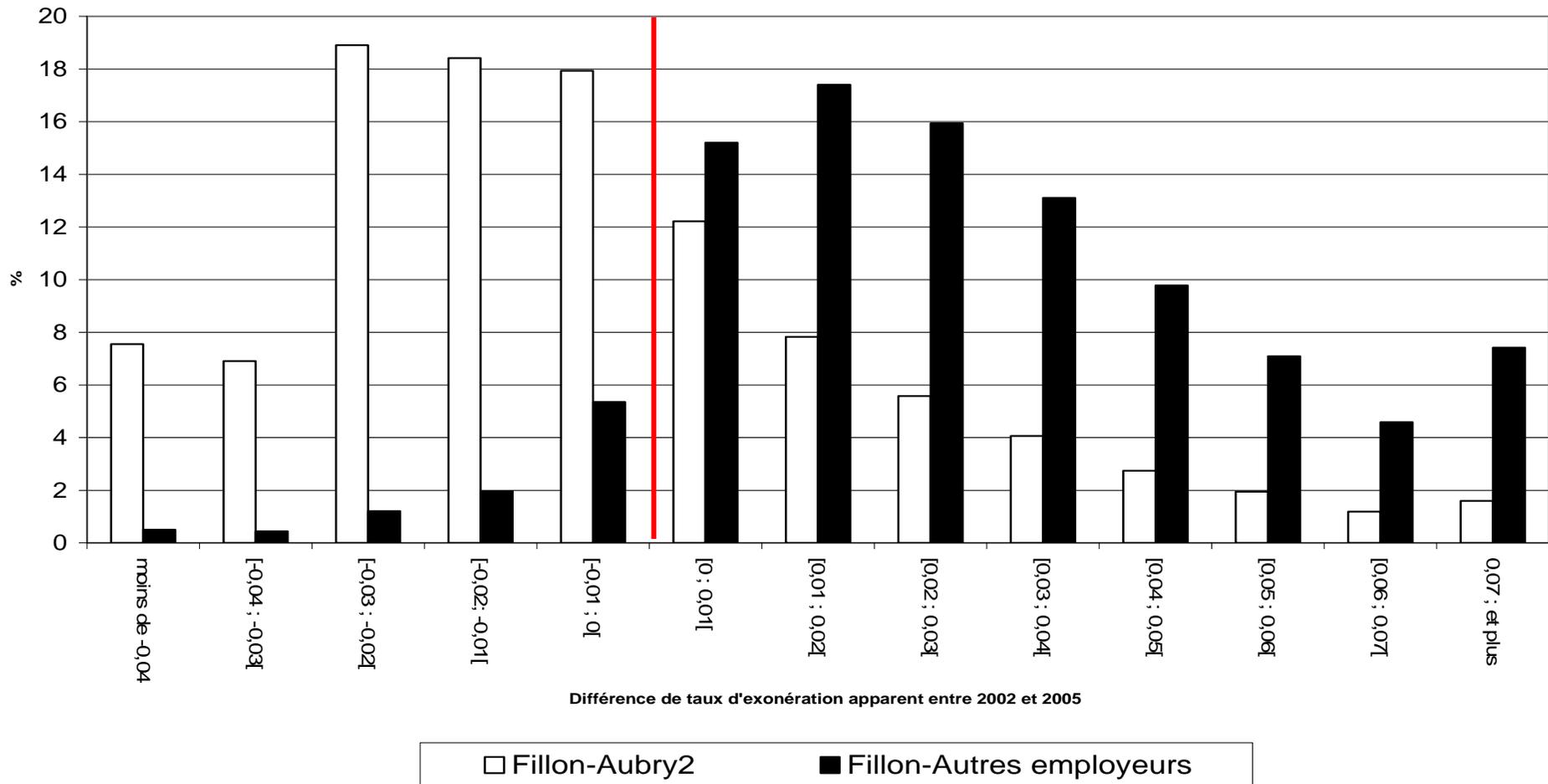
Calculé en 2005 avec correction

$$\frac{Cotexo(\widetilde{b}^{05}, \widetilde{dsal}^{02})}{Sal^{02} + PP(\widetilde{b}^{05}, \widetilde{dsal}^{02})}$$

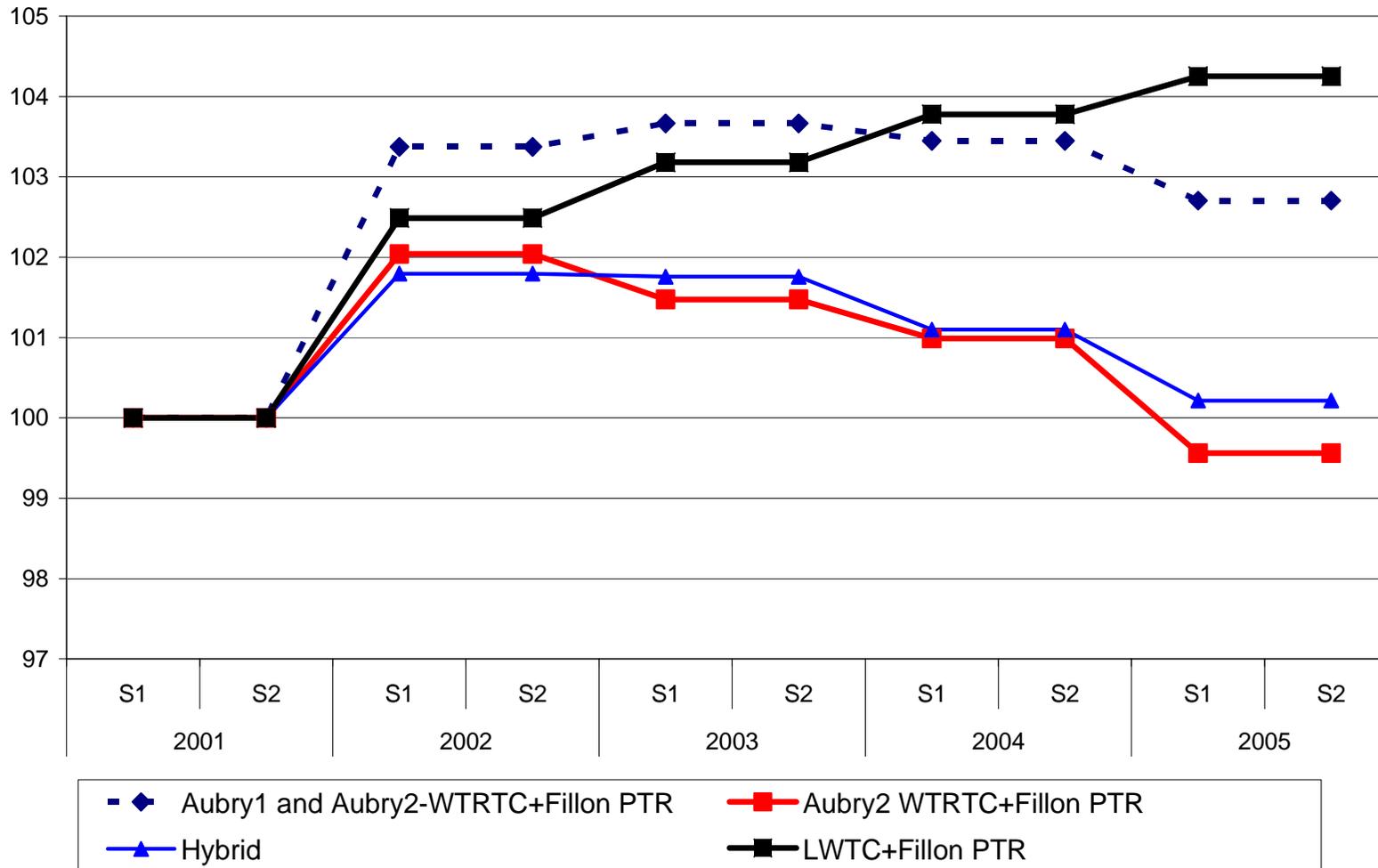
Calculé en 2005 avec correction
et en tenant compte des hausses du Smic

$$\frac{Cotexo(\widetilde{b}^{05}, \widetilde{dsal}^{02})}{\widetilde{Sal}^{02} + PP(\widetilde{b}^{05}, \widetilde{dsal}^{02})}$$

Anatomie du choc de cotisations (2)



Evolution de l'emploi par catégorie d'établissements



Resultats des estimations (extraits)

Elasticités au coût du travail pour les entreprises à 39 heures

	MCO	Regressions quantiles		
		25%	50%	75%
Effectifs occupés	-0.2234*** [-0.3414; -0.1054]	-0.2303*** [-0.3554; -0.1052]	-0.1512*** [-0.2439; -0.0585]	-0.1424** [-0.2740; -0.0107]
Equipalent temps complet	-0.3133*** [-0.4252; -0.2015]	-0.3828*** [-0.5181; -0.2476]	-0.1260*** [-0.2259; -0.0261]	-0.1680*** [-0.2980; -0.0379]
Salaire brut	-0.4484*** [-0.5082; -0.3885]	-0.6480*** [-0.7124; -0.5836]	-0.2774*** [-0.3340; -0.2208]	-0.1171*** [-0.1889; -0.0453]

Résultats (complets)

❑ **Exonérer est bon pour l'emploi**

Dans les deux types d'entreprises, 1 % de baisse du coût du travail supplémentaire donne environ 0,25 % à 0,30 % de hausse de l'emploi

❑ **Pour autant, la réforme Fillon a eu un effet globalement balancé sur l'emploi**

- Hausse de l'emploi dans les entreprises à 39 heures
- Baisse dans les entreprises à 35 heures (cf. anatomie du choc)

❑ **L'effet de la réforme sur l'emploi total n'est que faiblement positif.**

❑ **Exonérer est bon pour les salaires aussi**

- Dans les entreprises à 39 heures, où la réforme a entraîné une nette baisse du coût du travail, celle-ci est reliée à une hausse des salaires
- Dans les entreprises à 35 heures, les effets sur le coûts du travail sont moins nets et il en va de même pour la hausse des rémunérations
- Globalement, la réforme a plutôt contribué à soutenir la progression des salaires.

Conclusions

- ❑ Nos résultats confirment que l'emploi est bien sensible au coût du travail (pas d'exception française).
 - ❑ Les exonérations de cotisations sociales constituent dans l'absolu de puissants relais pour améliorer la situation de l'emploi.
 - ❑ Pour autant, la réforme Fillon, qui a harmonisé les exonérations en les redistribuant entre les entreprises (de celles à 35 heures vers celles à 39 heures) n'a eu que des effets globalement limités sur l'emploi.
- ⇒ Message de politique économique : « Sanctuariser » les allègements de charge en France.

En savoir plus...

Les effets des allègements de cotisations sociales sur l'emploi et les salaires : une évaluation de la réforme Fillon de 2003

Matthieu Bunel, Fabrice Gilles, Yannick L'Horty, *Document de travail du CEE*, n°122, août 2009.

Une évaluation des effets des baisses de cotisations sociales sur les bas salaires dans le cadre de la réforme Fillon de 2003

Matthieu Bunel, Richard Duhautois, Fabrice Gilles, Patrick Kwok, Yannick L'Horty, Marianne Pauchet, Corinne Perraudin (Juillet 2009), *Rapport de recherche du CEE*, n°55, juillet 2009.

