

Alain TRANNOY, Directeur d'études à l'EHESS, Professeur à l'Ecole d'Economie d'Aix-Marseille. Conseiller scientifique à France-Stratégie

Rapporteur référent sur l'évaluation réalisée par TEPP

Réactions sur les travaux du TEPP et de LIEPP

21 mars 2017

Les stratégies économétriques suivies par les deux équipes TEPP et LIEPP ont été différentes depuis le départ. Sur certains aspects, ces différences de stratégies n'ont pas porté à conséquence et les résultats sont convergents. C'est en particulier le cas de l'impact du CICE sur l'investissement (aucun impact causal différentiel selon le degré d'exposition au CICE, sur la recherche et développement (même résultat), sur les marges (résultat positif). En revanche, sur tout ce qui touche à l'emploi (ou à la structure de l'emploi), les résultats suivant les spécifications peuvent être négatifs, nuls ou positifs. Une telle fragilité des estimations à la spécification retenue révèle en soi que l'effet sur l'emploi a été les deux premières années de faible ampleur, proche d'un effet nul. Au jour d'aujourd'hui, il est difficile d'affirmer à coup sûr qu'il a été positif, même si la plausibilité penche du côté positif et nous allons expliquer pourquoi. Depuis 6 mois, les équipes, France Stratégie, et les experts tentent de mieux comprendre l'origine des différences de résultat. C'est une sorte de puzzle dont nous allons chercher à reconstituer toutes les pièces.

En remarque liminaire, nous pensons qu'il faut privilégier les résultats lorsque l'on pondère les entreprises par leurs poids d'échantillonnage. Nous nous intéressons avant tout à l'effet macroéconomique sur l'emploi, et il est bien évident que l'effet sur Renault ne peut avoir le même poids que l'effet sur une TPE. Le fait de ne retenir que les résultats pondérés a déjà pour effet de mettre de côté beaucoup d'effets négatifs. Ensuite, nous pensons qu'il faut privilégier les études en panel, qui permettent naturellement de tenir compte de l'hétérogénéité des entreprises qui est très importante, en termes de taille, de secteur, etc.

1. Le point d'entrée naturel de l'analyse économétrique nous semble devoir être un modèle en panel en niveau avec effet fixe entreprise et des contrôles. On estime alors un modèle en double différence entre les entreprises les plus exposées au CICE et les entreprises les moins exposées au CICE et on instrumente le traitement. C'est le choix suivi par le LIEPP. Estimer un modèle en différence première permet de ne pas avoir à estimer les effets fixes, c'est le choix suivi par TEPP mais a priori la différence de résultat ne peut venir de là. A partir de là, les stratégies suivies par les deux équipes diffèrent sur 3 points.

D'abord, l'échantillon d'estimation de TEPP est plus petit que celui du LIEPP et ne comprend que des entreprises de plus de 5 salariés. Le nettoyage du fichier d'entreprises peut aussi introduire des différences d'échantillonnage.

Ensuite, le LIEPP fait l'hypothèse que l'effet du traitement est linéaire, alors que TEPP introduit une possible de non-linéarité du traitement. TEPP autorise que l'effet du CICE soit plus ou moins important selon le degré d'exposition des entreprises, les entreprises étant réparties en quartiles, du quartile le moins exposé (le taux de subvention à la masse salariale est le plus faible) au quartile le plus exposé. Nous pensons là que la démarche de TEPP est plus à même de détecter un effet positif. En effet, il est assez connu que les agents économiques, qu'ils soient entreprises ou ménages, ne réagissent pas forcément à un changement très marginal des incitations. Il faut que le changement dépasse un certain seuil

pour provoquer un changement de comportement. Toutefois, il est naturel de penser que les entreprises les plus exposées soient très différentes des entreprises les moins exposées. Elles n'ont pas par définition la même proportion de cadres, elles n'évoluent pas dans le même secteur économique, etc. Dès lors, si on trouve une réaction des entreprises du quartile le plus exposé par rapport à celles appartenant au moins exposé, on peut se demander si cela n'est pas dû à la nature des entreprises. En d'autres termes, il n'est pas sûr que si les entreprises les moins exposées avaient été autant exposées que les plus exposées, elles auraient réagi de la même façon. On retrouve là une difficulté détectée depuis le départ dans l'évaluation ex post du CICE. Une condition pour une évaluation propre d'un dispositif commande que les agents économiques traités et non traités, ici peu traités, soient identiques en tout point, condition non vérifiée par nature avec le CICE.

La troisième différence tient dans le nombre de variables de contrôles introduits par les deux équipes. TEPP introduit beaucoup plus de groupes de contrôle que le LIEPP. Cela découle assez logiquement du fait que TEPP essaie plus d'aller capter la différence de réponse au CICE due au degré d'exposition du CICE. Ce faisant, comme expliqué plus haut, il capte aussi une différence de sensibilité au traitement due à la différence de nature des entreprises. Pour essayer de contrecarrer ce dernier effet, TEPP introduit une trentaine de contrôles, là où le LIEPP se contente de 5 contrôles. Mais, signalons que comme les estimations de TEPP sont réalisées en différence première, un certain nombre de contrôles sont aussi introduits en différence première.

Les travaux menés depuis 6 mois permettent de mieux comprendre l'origine des différences de résultat entre TEPP et LIEPP sans que l'élucidation soit complète.

Quand l'effet du traitement est linéarisé, ce qui signifie que l'on suppose que toutes les entreprises réagissent avec la même intensité à une variation du traitement (taux du CICE par rapport à la masse salariale), alors le LIEPP ne trouve aucun résultat, quelque soit le nombre de variables de contrôles, résultat confirmé par TEPP.

Quand on discrétise la variable de traitement et qu'on passe à un modèle non-linéaire, TEPP trouve un effet significatif dans le dernier quartile en 2013 et encore plus en 2014. Ce sont les nouveaux résultats de TEPP en panel estimés en différences premières avec plus de 30 contrôles équivalent à une estimation en panel en niveau, sauf que certaines variables de contrôles sont exprimées en différences premières et d'autres en niveau.

A ce stade, je privilégie les résultats de TEPP parce que je pense que la non-linéarité de l'effet du traitement selon le degré d'exposition est vraisemblable, comme expliqué plus haut, avec comme conséquence probable, malheureusement, que l'on capte aussi une différence due à la différence de nature des entreprises les plus exposées, un effet de sélection du traitement. Pour prendre une image dans le secteur médical, est ce que les patients les plus malades ont mieux réagi parce qu'on leur a administré plus de traitement, ou parce qu'ils étaient plus malades ? Malgré, les précautions prises par TEPP, il est difficile d'être affirmatif que c'est la première raison qui l'emporte.

2. L'histoire pourrait s'arrêter là. En effet, pour que la double différence soit un estimateur valide de l'effet causal, il faut vérifier une tendance de trend commun entre les groupes de traitement et de contrôle, test dit placebo. Le test placebo n'est vérifié ni pour LIEPP ni pour TEPP. C'est un coup dur pour la stratégie d'identification de l'effet causal suivie par les deux équipes, parce que cela invalide l'estimateur de double différence et donc les résultats trouvés dans l'étape 1.

3. Pourquoi l'échec du test placebo ? Les mesures prises en 2011 et 2012 d'alourdissement de la pression fiscale sur les entreprises (et entre autres sur le forfait social) subies plus particulièrement par celles ayant beaucoup de cadres pourraient en être la cause. Ces entreprises figurent dans le premier quartile du TEPP, le groupe ayant bénéficié le moins du CICE. Or, l'effet placebo est testé en 2012, là où se produit cet alourdissement différentiel de la taxation du travail. Cela pourrait expliquer que le quartile le plus traité et le quartile le moins traité n'aient pas le même trend d'emploi en 2012. Ensuite, on peut développer une argumentation comme quoi ces mesures auraient pour conséquence que l'effet positif obtenu par TEPP soit surestimé. Peut-être que l'ampleur du biais est marginale. Mais le sens du biais est plutôt à une surestimation de l'effet positif sur l'emploi.
4. Comme la double différence est invalidée, il est normal de tester une stratégie d'estimation en triple différences et d'introduire un trend temporel par entreprise. Mais un nouvel écueil se dresse sur le chemin des équipes, le manque de profondeur de la période d'estimation du modèle, 5 ans donc 4 différences premières d'emploi, voire 3 différences premières d'emploi, et là les deux équipes ne trouvent pas d'effet sur l'emploi.

Donc au total, on obtient un effet significatif positif sur l'emploi dans le dernier quartile avec l'estimation en double différence mais la double différence est invalidée, et de plus un biais de surestimation, peut-être de faible ampleur, est détecté. Et aucun effet sur l'emploi avec la triple différence mais ce modèle est mal estimé du fait de la trop faible durée de la période d'estimation. A ce stade, la conclusion à destination du grand public ne peut être à mon avis que si l'effet sur l'emploi a été positif, il a été en 2013 et 2014 de faible ampleur et il n'est pas encore totalement sûr à ce stade du travail scientifique qu'il ait été positif. En revanche, compte tenu de ce qu'on sait sur l'impact causal sur les allègements de charge, il est à peu près sûr que ses effets auraient été plus puissants. En revanche, on ne peut pas inférer des résultats très modestes sur l'emploi du CICE les deux premières années de montée en charge du dispositif qu'il en est allé de même pour les années ultérieures.

En ce qui concerne l'effet sur l'emploi différencié par catégorie de travailleurs, seuls les résultats de TEPP sont en accord avec la rationalité économique standard des entreprises, avec une augmentation de la part des ouvriers et employés au détriment des cadres. Quand une entreprise embauche ou conserve un salarié, elle réalise toujours une analyse coût-bénéfice. Lorsqu'une entreprise minimise son coût, et qu'elle peut substituer dans une certaine mesure des travailleurs d'un certain degré de qualification à d'autres d'un degré de qualification plus ou moins élevé, elle compare le rapport des coûts du travail au rapport des productivités marginales. En deçà de 2,5 SMIC, le rapport des coûts du travail est resté le même, puisque le pourcentage de réduction de la masse salariale du CICE est constant, 6% en 2014. En conséquence, le CICE n'introduit aucune incitation à modifier le mixte de qualification en deçà de 2,5 SMIC. En revanche, le coût d'un salarié en deçà de 2,5 SMIC a diminué par rapport au coût du travail au-delà de 2,5 SMIC. Nous concluons que les entreprises ont pu arbitrer en faveur de relativement plus d'embauches des premiers et en défaveur des seconds mais pas dans le sens contraire, selon la théorie standard. Si les résultats empiriques consolidés devaient contredire ce raisonnement économique de base, ils devraient être étayés par une théorie spécifique.