



Analyse critique du document

Trends in Income Inequality and
its Impact on Economic Growth
(Federico Cingano)

OECD Social Employment and Migration Working Papers, no 163, 2014

par la

Société de Calcul Mathématique SA

Rédaction: Miriam Basso et Bernard Beauzamy

Février 2015

Résumé opérationnel

Cet article est complètement dépourvu de contenu scientifique. Il prend en entrée un certain nombre de quantités non mesurables, leur affecte arbitrairement une valeur numérique, les relie par une formule entièrement artificielle, et prétend en déduire une influence des inégalités de revenu sur la croissance.

Pour prendre une comparaison, illustrant à merveille la qualité de la méthodologie, on peut vouloir noter un restaurant, à partir de critères comme la présentation des plats, la forme de l'assiette, l'intensité de l'éclairage, le prix des restaurants aux alentours. On mélangera tous ces critères au moyen d'une formule mathématique appropriée et on conclura que les restaurants chinois sont les meilleurs, et que le facteur qui influe le plus sur la note est l'origine ethnique de l'aide cuisinier : tout ceci s'obtient très simplement, si l'on fait les bons réglages dans la formule.

Les mathématiques ne sont utilisées ici que pour donner une apparence scientifique à la démarche. L'auteur les maîtrise mal, et d'innombrables erreurs sont commises : nous les passons en revue plus loin.

Du point de vue de l'aide à la décision (appui aux politiques publiques), on note dès l'abord une profonde malhonnêteté ; dans l'introduction, l'auteur écrit : *"S'appuyant sur des données harmonisées couvrant les pays de l'OCDE au cours des 30 dernières années, l'analyse économétrique suggère que les inégalités de revenus ont un impact négatif et statistiquement significatif sur la croissance ultérieure. En particulier, ce qui importe le plus est l'écart entre les ménages à faible revenu et le reste de la population."*

Il constate (ou plutôt croit constater, moyennant toutes les incohérences que nous décrivons plus loin) une corrélation entre deux indicateurs : les inégalités et la croissance. Cette conclusion est entièrement artificielle et ne résulte que des "bricolages" faits par l'auteur. La présenter comme un lien de causalité ("impact significatif") est une malhonnêteté.

Prenons une comparaison pour bien faire comprendre ceci : lorsqu'il pleut, les riches sortent leur voiture et les pauvres sortent leur imperméable. Taxer les riches, taxer les voitures, n'empêchera pas la pluie de tomber et les pauvres d'avoir besoin d'un imperméable.

English Summary

This paper has absolutely no scientific contents. It takes, as input data, some non measurable quantities, attributes to them an arbitrary numerical value, links them with a completely artificial formula, and pretends that it demonstrates an influence of income inequalities upon economical growth.

Just to make a comparison, which clearly shows the quality of the methodology, one may give a grade to any restaurant, using criteria such as the presentation of the dishes, the shape of the plates, the intensity of the light, the price of the restaurants in the neighborhood. One would mix all these criteria using some appropriate mathematical formula and one would conclude that Chinese restaurants are the best, and that the factor which has the strongest influence upon the ranking is the ethnical origin of the cook help. Any such result is simple to obtain : just put the proper coefficients and weights in the formula.

Mathematics is used here only in order to give a scientific appearance to the whole paper. The author is not familiar with statistics and makes a large number of mistakes: we will indicate them later.

Such a paper is supposed to be used for decision help (public policies). So, it should be honest. However, from the very beginning, we see clearly that this is not the case, since the author writes in his Abstract : *"Drawing on harmonised data covering the OECD countries over the past 30 years, the econometric analysis suggests that income inequality has a negative and statistically significant impact on subsequent growth. In particular, what matters most is the gap between low income households and the rest of the population."*

The author sees (or, more precisely, pretends he sees, taking into account all logical mistakes we indicate later), some correlation between two indicators: inequalities and economical growth. Such a conclusion is totally artificial and is simply a result of the "bricolage" (that is, non-professional work of poor quality) made by the author. To present it as a causality link (*"significant impact"*) is completely dishonest.

Let us illustrate this with a simple and clear comparison: when it rains, rich people take their cars and poor people take their raincoats. If you put taxes on the rich, on the cars, this will not prevent the rain from falling, and will not prevent the poor from taking their raincoats.

I. Introduction

Ce rapport présente une analyse méthodologique de l'article "Trends in Income Inequality and its Impact on Economic Growth", de Federico Cingano, OECD Social Employment and Migration Working Papers, no 163, 2014. Il est rédigé par la Société de Calcul Mathématique SA.

II. Analyse de l'article

L'article de F. Cingano présente une analyse de l'impact de l'évolution des inégalités de revenus sur la croissance économique pour l'ensemble des pays de l'OCDE. Il se structure comme suit :

1. *Evolution historique des inégalités*

Les données OCDE utilisées montrent que les inégalités de revenus entre les pauvres et les riches n'ont cessé d'augmenter au cours des 30 dernières années pour la plupart des pays concernés, aussi bien avant qu'après la période de récession (2007).

Mais il y a déjà là un point très discutable. F. Cingano prend pour exemples les 10% plus bas et les 10% plus hauts (en termes de revenus) et non la répartition complète des revenus. Vouloir mettre en évidence l'évolution des disparités de revenus en se limitant aux deux quantiles extrêmes constitue certainement une erreur méthodologique, puisqu'on ignore ainsi l'évolution de 80 % de la population, même si l'indicateur de Gini prend en compte un cumul de pourcentages.

Les 10% ayant les plus bas revenus ne sont pas directement liés à la croissance économique du pays ; ils relèvent d'une logique propre : il y a parmi eux un certain nombre d'inadaptés sociaux, d'immigrés, etc.

Les 10% ayant les plus hauts revenus ne sont pas non plus directement liés à la croissance économique du pays et relèvent également d'une logique différente. Il y a parmi eux un certain nombre de personnes vivant simplement de leurs revenus.

On constate donc cette double erreur de logique dans l'exemple pris par F. Cingano : d'une part, il ignore 80% de la population, et d'autre part les 20% qu'il retient relèvent de critères singuliers, propres à leur situation.

Par ailleurs, la comparaison des évolutions d'inégalité est présentée pour une période très étendue et des pays très différents. Nous pouvons nous attendre à ce que la qualité de l'information soit très hétérogène. L'article affirme que les données présentées ont été "harmonisées", toutefois ce processus d'harmonisation n'est pas présenté.

Les imperfections bien connues des indices de Gini (ils ne font pas la différence entre inégalités de bas revenus et inégalités de hauts revenus) ne sont pas prises en compte.

2. Bilan des études précédentes

La recherche bibliographique menée par l'auteur montre qu'il n'y a pas de consensus sur le lien existant entre les inégalités de revenus et la croissance économique. Parmi les éléments signalés comme problématiques, sont cités :

- qualité des données, souvent qualifiée d'insuffisante ;
- indicateurs choisis : leur pertinence est souvent remise en cause ;
- méthodes utilisées : elles sont jugées inappropriées car reposant sur trop d'hypothèses factices.

Mais cela n'empêche manifestement pas F. Cingano de prétendre en introduction que lui, grâce à ses méthodes, parvient à une conclusion claire. Les réserves faites ne sont nullement reprises dans le résumé à l'usage des décideurs. Ceci est une contradiction claire avec l'une des règles fondamentales de la recherche scientifique : la conclusion doit mentionner tous les doutes et ne pas être orientée.

3. L'approche mathématique utilisée

L'approche utilisée par F. Cingano est la suivante :

Une équation empirique est donnée où la croissance est estimée comme une fonction linéaire de l'inégalité; elle est donnée sous la forme :

$$\ln(y_{i,t}) - \ln(y_{i,t-1}) = \alpha \ln(y_{i,t-1}) + X_{i,t-1}\beta + \gamma Ineq_{i,t-1} + \mu_i + \mu_t + \epsilon_{i,t}$$

Où :

- i est l'indice du pays ;
- $(t, t-1)$ correspond à une période de 5 ans ;
- $\ln(y)$ est le logarithme népérien du PIB réel par habitant ;
- $Ineq$ représente l'inégalité de revenus ;
- le vecteur X contient l'information inhérente au capital humain et physique ;
- μ_i et μ_t sont calculés à partir d'un échantillon de données propre à chaque pays ;

La formule ci-dessus, totalement empirique, est absolument dépourvue de signification. Pourquoi utiliser le logarithme du PIB, plutôt que sa racine carrée ou le cosinus ? Elle a une apparence mathématique, propre à impressionner le lecteur non informé, mais ne repose sur aucun fondement.

L'auteur dit que cette équation a été retenue suite aux études réalisés par Sollow. Toutefois, les recherches bibliographiques effectuées indiquent que "aucun modèle économétrique probant ne peut être retenu pour répondre à la question". Dès lors, pourquoi reprendre cette formule ?

A l'intérieur de la formule elle-même, les grandeurs utilisées ne peuvent être correctement définies. Passons sur le PIB, dont on connaît les imprécisions. Mais quel sens numérique donner au "capital humain" ? Nous en parlons plus loin.

Les choix portant sur les coefficients sont arbitraires ; on ne sait pas par quel moyen, au moyen de quelles données, ils ont été calculés.

Aucune analyse de sensibilité n'est réalisée. Aucune incertitude sur les données n'est prise en compte.

L'article prétend conclure que l'éducation joue un rôle (ce que chacun est enclin à admettre). Mais l'éducation n'est décrite qu'au travers d'un paramètre à trois valeurs (faible, moyen, fort), dont la valeur est discutable.

Le "capital humain" est lui-même estimé à partir d'une formule empirique :

$$HC_{i,t,c} = \beta_1 PEB_{i,t,c} \times Ineq_{i,t,c} + \beta_2 PEB_{i,t,c} + \theta X_{i,t,c} + \mu_t + \mu_c + \varepsilon_{i,t,c}$$

où, en particulier, $PEB_{i,t,c}$ est un ensemble de trois indicateurs reflétant le niveau d'éducation des parents, etc.

Il est complètement évident que, à partir d'une telle formule, entièrement dépourvue de contenu et de signification, on peut donner au capital humain n'importe quelle valeur souhaitée : l'approche relève ici de la caricature.

On note au passage que l'auteur (est-ce malhonnêteté ? est-ce incompetence ?) fait intervenir les inégalités deux fois : d'une part à l'intérieur de la formule principale, et encore une fois à l'intérieur de la définition du capital humain. Il n'est pas étonnant, dans ces conditions, qu'il parvienne au résultat selon lequel les inégalités jouent un rôle important en sortie, puisqu'il s'est arrangé pour les mettre en valeur en entrée.

III. Conclusions

A. Sur l'aspect économique

Il n'est pas possible de réaliser un modèle mathématique, quel qu'il soit, incluant des concepts comme le "capital humain" ou l'éducation, car ils sont fondamentalement non-numériques. On peut évidemment parler d'un niveau de formation ou d'éducation, dé-

terminer par exemple le pourcentage de population, dans chaque pays, qui a le bac, ou qui poursuit ses études à Bac+2, Bac+5, ou ce que l'on voudra. Mais ces indicateurs instantanés n'ont que peu de valeur explicative ou comparative : le niveau du baccalauréat et ce qu'il représente varient d'une époque à l'autre, d'un pays à l'autre.

En France, par exemple, le pourcentage de jeunes ayant réussi le baccalauréat a considérablement progressé depuis la seconde guerre mondiale, sans que l'on puisse dire pour autant que les jeunes soient mieux formés ou que le pays se porte mieux.

B. Rappel des règles fondamentales de la recherche scientifique

L'article ici présenté déroge de manière essentielle aux règles fondamentales de la recherche scientifique :

- Les données d'entrée doivent toujours être rapportées avec leurs incertitudes ;
- Une analyse de sensibilité doit toujours être faite pour tenir compte de ces incertitudes. Par exemple, la conclusion est-elle maintenue si les données d'entrée sont augmentées de 10% ou diminuées de 20 % ? Cette question n'est même pas posée ici.
- Un modèle, quel qu'il soit, doit toujours être validé sur d'autres données que celles qui ont permis de le constituer. L'auteur aurait pu constituer une formule avec certains pays, ou certaines époques, et le tester ensuite sur d'autres pays ou d'autres époques. Il n'y a ici aucune tentative en ce sens.
- Le résultat présenté doit impérativement inclure toutes les réserves, tous les doutes, qui ont pu apparaître pendant l'investigation. Ici c'est tout l'inverse : le résumé ne reprend aucune des difficultés méthodologiques connues de l'auteur.

C. Une présentation malhonnête

Il y a une confusion très grave, et certainement malhonnête, entre constater une corrélation et annoncer qu'il y a un lien de cause à effet. Annoncer, comme l'auteur le fait dans son résumé à l'usage des décideurs "*les inégalités de revenus ont un impact négatif et statistiquement significatif sur la croissance ultérieure*" est une conclusion très forte, résultant du mot "impact significatif". Or il n'a absolument rien montré de cela. Il croit avoir montré que là où l'on observe des inégalités fortes, la croissance est faible : ce n'est pas du tout la même chose. La conclusion de l'auteur veut amener les politiques à taxer davantage les hauts revenus : c'est clairement le but affiché. Mais il n'y a rien dans la démarche logique qui permette de penser que taxer les hauts revenus engendrera la croissance.