



UNIVERSITÉ DE NANTES



IAE NANTES
ÉCONOMIE & MANAGEMENT

Les déterminants du taux de chômage en France entre 1982 et 2017

Université de Nantes 2019

DAVOINE Nicolas

AUROUET Lucas

SOMMAIRE

I - Introduction

II - Présentation du sujet

II. 1. -Mutation du marché du travail en France depuis les années 1980

II. 2. - La mutation de la nature de l'emploi

II. 3. - Le coût du travail élevé

II. 4.- L'historique des politiques de l'emploi en France

II. 5. - La réduction du coût du travail

II. 6. - Justification théorique des variables explicatives

III - Modèle théorique RLM

III. 1. - Les hypothèses du modèle RLM

III. 2. - Caractéristiques des estimateurs ponctuels

III. 3. - Estimateurs par intervalle de confiance

IV. - Analyse économétrique du sujet

IV. 1. - Définitions des variables

IV. 2. - Modèles intermédiaires

IV. 3. - Modèles définitif

V - Conclusion

Les déterminants du taux de chômage en France entre 1982 et 2017

I - Introduction

Nous avons, dans ce dossier, choisi de nous intéresser aux déterminants du chômage en France entre 1982 et 2017. Le chômage, est l'une des variables macroéconomiques essentielles pour un gouvernement. Sa mesure s'effectue à l'aide du taux de chômage qui se compose du ratio population active inoccupée/ population active totale. Sa définition (i.e. la définition de « inoccupée ») varie selon les organismes de mesure, on en retient cependant deux qui sont les plus répandues, celle du Bureau International du Travail (BIT) et celle de Pôle Emploi.

Définition du BIT : *un chômeur est une personne en âge de travailler (15 ans ou plus) qui répond simultanément à trois conditions : être sans emploi (c'est à dire ne pas avoir travaillé au moins une heure durant une semaine de référence), être disponible pour prendre un emploi dans les 15 jours, avoir cherché activement un emploi dans le mois précédent ou en avoir trouvé un qui commence dans moins de trois mois.*¹

Définition de Pôle Emploi : *Le chômeur ne pas avoir d'emploi ou disposer d'une activité réduite, de moins de 78heures pour le mois, le chômeur doit être disponible immédiatement, le chômeur doit rechercher un emploi à plein-temps, puisqu'il peut exercer une activité réduite, le chômeur doit être inscrit à Pôle Emploi.*²

Le taux de chômage et ses déterminants sont au cœur de l'analyse économique. Les controverses à ce sujet sont nombreuses, beaucoup d'économistes sont en désaccord sur les variables qui peuvent l'influencer. Selon les hypothèses prises en compte dans les modèles économiques qui traitent la question, les résultats sont souvent contradictoires. C'est pourquoi il est primordial de vérifier empiriquement les mécanismes qui sont suggérés dans ces modèles. C'est ce à quoi nous nous emploierons dans ce dossier.

1 Organisation internationale du Travail, définition du Bureau International du Travail : <https://www.ilo.org/global/about-the-ilo/who-we-are/international-labour-office/lang--fr/index.htm> consulté le 25/03/2019.

2 <http://statistiques.pole-emploi.org/stmt/publication> consulté le 25/03/2019.

II- Présentation du sujet

II. 1. -Mutation du marché du travail en France depuis les années 1980

(Tableau1)

1. Effectifs et taux d'activité pour l'ensemble des personnes de 15 à 64 ans

	1975	1980	1985	1990	1995	2000	2005	2012
	en milliers							
Population active	22 082	23 500	24 208	24 710	25 246	26 145	27 267	28 328
Emploi à temps plein	19 703	20 438	19 631	19 985	19 132	19 820	20 594	21 013
Emploi à temps partiel	1 613	1 786	2 414	2 751	3 568	4 086	4 241	4 507
Chômage	766	1 275	2 163	1 975	2 547	2 239	2 431	2 808
Population	32 234	33 572	35 820	36 827	37 193	37 866	39 011	39 950
Taux d'activité (en %)	68,5	70,0	67,6	67,1	67,9	69,0	69,9	70,9

Champ : France métropolitaine, population des ménages, personnes de 15 à 64 ans.

Note : données corrigées pour les ruptures de série.

Source : Insee, enquêtes Emploi.

En 30 ans le marché du travail a subi des mutations légales pour concorder avec la conjoncture. Il a aussi été impacté par les évolutions démographiques (vieillessement de la population, migration,...). La population active (au sens du BIT) est passée de 22 millions à 28,3 millions entre 1975 et 2012, cette évolution se décompose en : 1,5 millions de nouveaux emplois à temps plein, 2 millions d'emplois à temps partiel et 1,2 millions de chômeurs. On peut premièrement remarquer que l'emploi à temps partiel a plus augmenté que l'emploi à temps plein. De plus le taux d'activité augmente de 2,4 points sur la période. Quels sont les causes de ces transformations ? Comment s'explique le taux de chômage et ses évolutions ?

En fait, le chômage explose en France à partir du milieu des années 1970, en période de post-premier choc pétrolier (1973) qui a bouleversé le monde économique, caractérisé en France par un ralentissement conjoint de la croissance et de la productivité, après trente années dites «Glorieuses». C'est à partir de cette période que le chômage dit «structurel» s'installe en France: il s'agit d'une partie du chômage qui s'est accumulé, que la France n'arrive pas à résorber... même en période de croissance soutenue. Le taux de chômage structurel (appelé aussi taux de chômage de long terme ou taux de chômage d'équilibre) est celui qui demeure, malgré une conjoncture économique favorable.³

3 Barou.V, Buridant.J, Figliuzzi.A, Montoussé.M, 2016, *Histoire des faits économiques*, Partie 4: 1970-2015 Le temps des ruptures p 277-278, 2016

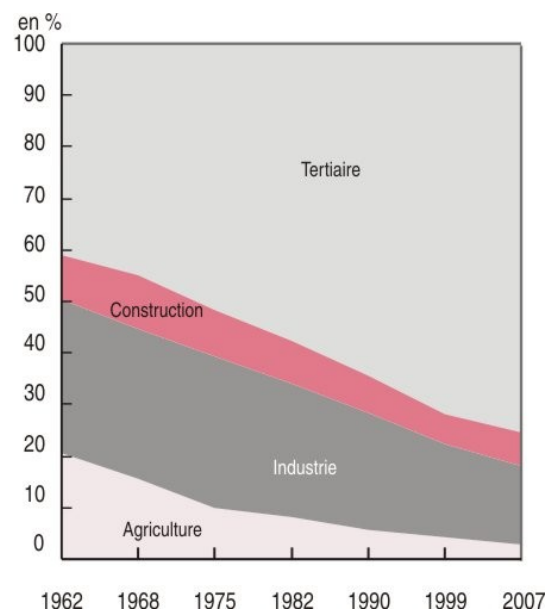
Le chômage structurel en France est estimé à environ 7%, ce niveau élevé révèle que le chômage structurel de masse constitue une spécificité nationale qui distingue la France des autres nations. Ainsi les politiques conjoncturelles ne peuvent avoir qu'un effet limité sur la baisse du chômage puisqu'elles n'influencent pas le chômage structurel. Cela suggère que les améliorations du marché du travail en France seront avant tout le fruit de réformes structurelles du fonctionnement global de l'économie (sur l'architecture des prélèvements et des transferts sociaux, politiques de l'emploi, de la concurrence, de l'éducation, de la formation professionnelle...).⁴

II. 2. - La mutation de la nature de l'emploi

II. 2. A. - La tertiarisation de la production

Le ralentissement de la croissance et l'avènement du chômage de masse ne sont pas les seuls bouleversements du monde du travail en cinquante ans. La structure de l'économie a beaucoup évolué. Ainsi, l'agriculture employait-elle 15% des actifs en 1968, contre moins de 3% actuellement. La part de l'industrie a aussi décliné, alors que celle des services (le secteur tertiaire) représente désormais plus des trois quarts des emplois, contre moins de la moitié il y a cinquante ans comme le montre le graphique ci-dessous.⁵

(Graphique 1)⁶



4 <http://dessinemoileco.com/dou-vient-cho%CC%82mage-y-t-il-remede/> consulté le 25/03/2019

5 .Barou.V, Buridant.J, Figliuzzi.A, Montoussé.M, 2016, *Histoire des faits économiques*, Partie 5: 1945-1973 Croissance et nouvel ordre économique international p 144, 2016

6 <https://www.insee.fr/fr/statistiques/1283207> consulté le 25/03/2019

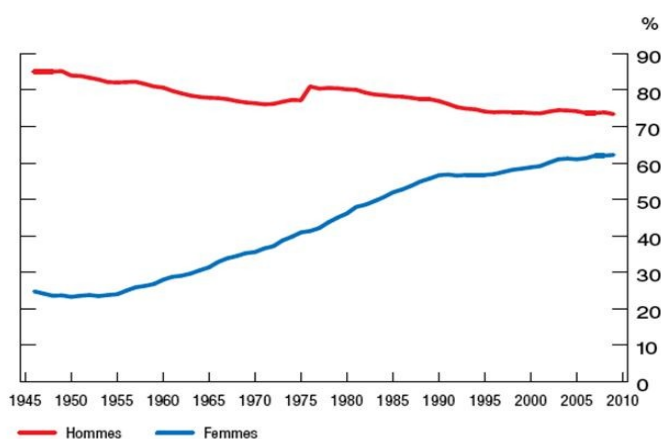
Ces changements s'accompagnent aussi d'une évolution du type d'emplois proposés, qui sont de plus en plus qualifiés. Le nombre de cadres ou de professions intermédiaires (agents de maîtrise, techniciens, commerciaux...) a explosé, alors qu'on compte de moins en moins d'ouvriers et d'agriculteurs.

II. 2. B. - Les femmes de plus en plus présentes sur le marché du travail

La tertiarisation de l'économie a aussi un effet indirect, il induit une féminisation de l'emploi ces 30 dernières années. Les femmes bénéficient d'un meilleur accès au marché du travail. Ainsi depuis les années 80 on recense 3,2 millions de femmes qui sont entrées sur le marché du travail (contre seulement +0,2 million pour les hommes). Les femmes occupent désormais 49 % des emplois en France alors qu'elles en occupaient seulement 41 % en 1980.⁷ Ces femmes sont surtout employées dans le secteur des services (aide à la personne, services administratifs, comptabilité). Les femmes sont de plus en plus nombreuses sur le marché du travail français, qui doit s'adapter, ce qui participe entre autres à l'augmentation du nombre de contrats à temps partiels.

(Graphique 2)⁸

Taux d'activité selon le sexe : 1945-2010



Nota : En 1976, la population active est redéfinie : désormais, au lieu de désigner les personnes de 14 ans et plus, celle-ci désigne les personnes âgées de 15 ans et plus.

Banque du Canada: 75 ans d'histoire en chiffres

7 <https://slideplayer.fr/slide/11357242/> consulté le 25/03/2019

8 <https://slideplayer.fr/slide/11357242/> consulté le 25/03/2019

II. 2. C. - Le développement du temps partiel

Le nombre de contrats à temps partiel s'est accéléré depuis le début des années 80. Plusieurs facteurs peuvent expliquer ce phénomène. Tout d'abord une plus grande fluctuation de la demande des ménages contraint les entreprises à adapter rapidement leur production, les incitant donc à embaucher de plus en plus d'employés pendant une courte période.

De plus la succession des périodes de récession depuis les années 1970 ne permet pas aux entreprises françaises de faire des anticipations sûres sur l'avenir, les incitant donc à se tourner vers des contrats courts plutôt que des contrats à durées indéterminées. En effet les entreprises cherchent à ajuster leur force de travail en fonction de la conjoncture, la législation plus souple leur permet d'employer des personnes en contrats partiels sur des périodes très limitées pour adapter très rapidement le potentiel de production à la demande anticipée⁹. En moyenne, en 2017, 1,2 million de salariés étaient employés en contrat court (CDD de moins de 3 mois ou Intérim) soit 4,5 % des emplois totaux. La part des contrats courts en France est passée de 1,0 % en 1982 à 4,4 % en 2007 puis s'est stabilisée autour des 4 % depuis¹⁰. Ils sont surtout utilisés dans le secteur de l'industrie pour embaucher de la main d'œuvre peu qualifiée (représente environ 22 % de la totalité des contrats courts signés). Avec le développement du temps partiel se développe aussi le sous-emploi c'est-à-dire une situation dans laquelle un individu alterne entre des périodes d'emploi et de chômage. Ces transformations qu'a subi le marché du travail français durant cette période sont les conséquences des politiques de l'emploi, des frontières de plus en plus floues entre l'inactivité, l'emploi, et le chômage qui représente ce que FREYSSINET appelle l'«halo» autour du chômage.¹¹

II. 3. - Le coût du travail élevé

La France est l'un des pays du monde où le salaire minimum est le plus élevé, mais également où le coût du travail auquel on retire le salaire versé par l'employeur est le plus important, ces deux caractéristiques du marché du travail français sont les principales causes du

9 Barou.V, Buridant.J, Figliuzzi.A, Montoussé.M, *Histoire des faits économiques*, Partie 4: 1970-2015 Le temps des ruptures p 278-279, 2016

10 <https://books.openedition.org/cirac/294?lang=fr#ftn1> « Regard historique sur le marché du travail en France »Heye.E, 2011 consulté le 25/03/2019

11 Fessiney. J, *Le Temps de travail en miettes : vingt ans de politique de l'emploi et de négociation collective*, Éditions de l'Atelier, Paris, 1997

coût du facteur travail élevé pour les entreprises nuisant à la flexibilité des salaires à cause du caractère structurel du chômage français dénoncé par la théorie néoclassique¹².

Tout d'abord, le Salaire Minimum Interprofessionnel de Croissance (SMIC) existe en France depuis 1950 et son augmentation correspond à l'inflation constatée l'année précédente sur la base de l'évolution de l'indice des prix à la consommation, plus 50 % de l'augmentation du pouvoir d'achat du salaire horaire de base ouvrier.

Le niveau élevé du salaire minimum est une particularité française. Bien qu'il existe dans beaucoup d'autres pays sous plusieurs formes différentes, c'est en France qu'il est le plus critiqué à cause de son niveau jugé trop élevé. Beaucoup jugent que le coût du travail pénalise trop les entreprises françaises qui embauchent pour des emplois peu qualifiés.

(Graphique 3)¹³



En plus d'un salaire minimum élevé, les entreprises françaises doivent payer des cotisations sociales importantes. Ces charges patronales qui sont les charges payées par les employeurs dans le cadre de la rémunération de leurs salariés, permettent de financer la sécurité sociale (assurance vieillesse, chômage et maladie, allocations familiales...).

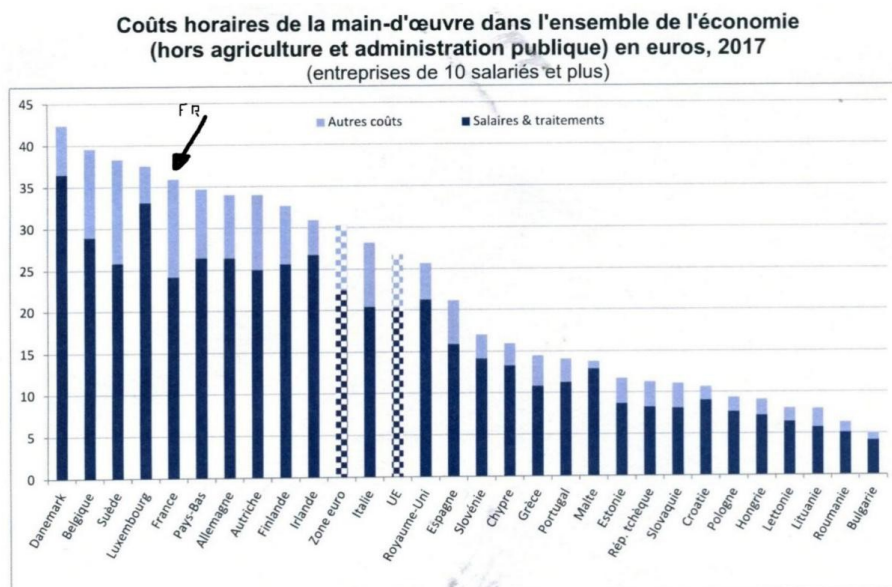
12 Pigou.A, *The theory of Unemployment*, Mac Millan, 1933

Rueff.J, *L'assurance chômage cause du chômage permanent*, Revue d'économie politique, 1931

13 <https://fr.statista.com/infographie/8345/salaire-minimum-les-pays-les-plus-generaux-au-monde/> consulté le 25/03/2019

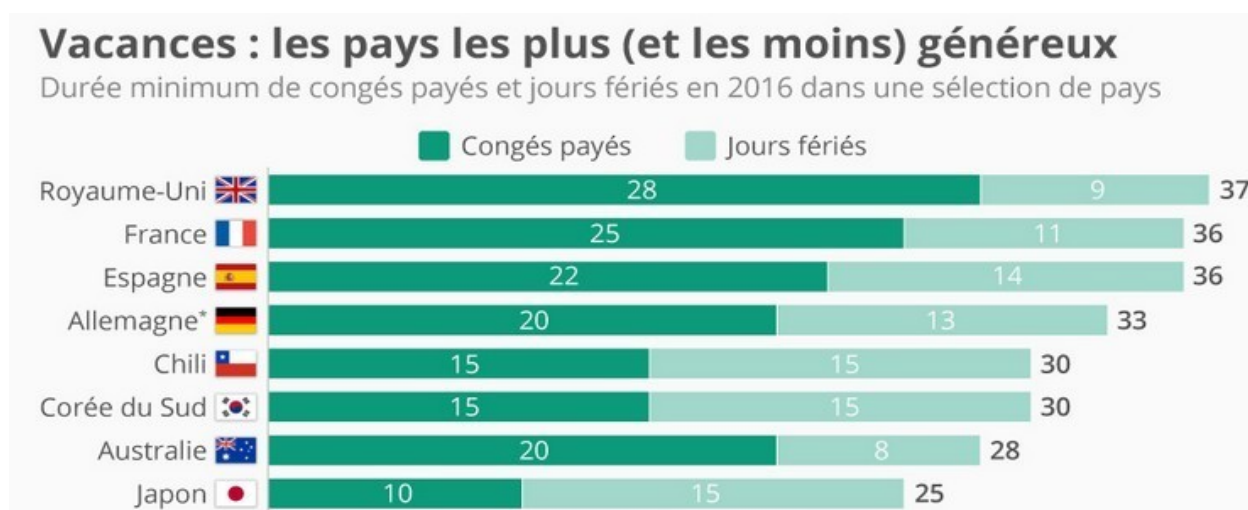
En France, aujourd'hui, les charges patronales représentent entre 25 et 42% du salaire brut d'un employés ce qui relativement élevé si on les compare aux autres pays d'Europe, comme le montre le graphique ci-dessous.

(Graphique 4)¹⁴



Par exemple, la France est un des pays où la loi offre le plus de congés payés aux salariés comme le montre le graphique ci-dessous:

(Graphique 5)¹⁵



14 <http://dominiegambier.fr/2018/04/les-couts-horaires-de-la-main-doeuvre-en-europe-entre-49e-et-425-e/> consulté le 01/04/2019

15 <https://fr.statista.com/infographie/15007/nombre-jours-feries-conges-payes-par-pays/> consulté le 01/04/2019

Ce coût du travail élevé est vivement critiqué par de nombreux économistes et hommes politiques, puisqu'il est l'une des principales causes de la faible compétitivité des entreprises françaises qui dégrade la balance commerciale, déficitaire depuis 2004. En 2018, les échanges commerciaux de marchandises de la France enregistrèrent un déficit de 59,9 milliards d'euros.¹⁶

II. 4.- L'historique des politiques de l'emploi en France

II. 4. A. - De 1950 aux années 1980, les politiques passives d'emploi

En France, les politiques de l'emploi ont principalement été des politiques passives d'emploi qui visent à indemniser ou à subventionner les chômeurs ou les personnes sorties du marché du travail, comme par exemple les politiques de pré-retraites qui permettaient de subventionner une période d'inactivité entre la cessation d'activité et la retraite, dans une optique de partage du temps de travail. De manière générale ces politiques permettent de garantir une stabilité des revenus, pendant une durée limitée, tout en protégeant contre l'exclusion sociale.

Cependant aux vues de la hausse du chômage dans les années 1980 et 1990 et de la dette publique, il est devenu clair que les politiques passives d'emplois coûtaient trop cher à l'Etat alors qu'elles ne permettaient pas de résorber le chômage, au contraire elles contribueraient à désinciter les chômeurs à chercher du travail étant moins rémunérant que les allocations chômage¹⁷.

II. 4. A. - Le passage aux politiques actives d'emploi

Stimuler l'offre de travail

A partir des années 1990 l'Etat français se tourne définitivement vers les politiques actives d'emploi qui visent à faire sortir rapidement les chômeurs du chômage en les incitant financièrement à chercher de manière active du travail, par exemple en les accompagnant davantage dans la recherche d'emplois ou en leur proposant des services de formation.

Les politiques actives d'emploi s'appuient sur la réduction du coût d'opportunité du travail, avec par exemple la mise en place du Revenu de Solidarité Active mis en place en 2008 afin de remplacer le Revenu Minimum d'Insertion, qui n'incitait pas assez le retour à l'activité. Le RSA

16 <http://www.lefigaro.fr/conjoncture/2019/02/07/20002-20190207ARTFIG00038-le-deficit-commercial-francais-a-continue-de-se-degrader-en-2018.php> consulté le 21/04/2019

17 Rosanvallon.P, *La crise de l'Etat-providence*, Seuil,1981, consulté le 21/04/2019

permet à la fois d'augmenter le revenu des travailleurs pauvres et d'encourager la recherche d'emploi puisque c'est un complément de revenu pour des individus travaillant à temps partiel ou au SMIC¹⁸.

La mise en place de la Prime Pour l'Emploi et plus tard de la Prime d'activité en 2015 pour remplacer les allocations précédentes s'intègrent dans une optique de réduction des revenus de remplacement sans conditions de recherche active de travail et de subventionner directement les individus lorsque la rémunération de l'emploi est trop proche des revenus de transfert¹⁹.

Stimuler la demande de travail

Ces types de mesures ont donc pour objectif de stimuler l'offre de travail, cependant le gouvernement a également fait en sorte, principalement depuis les années 1990, d'augmenter la demande de travail des entreprises grâce notamment aux emplois aidés. Les emplois aidés sont des emplois où l'Etat prend en charge une partie des cotisations patronales comme par exemple les Contrats Emploi-Solidarité, les Emplois d'Avenir pour les jeunes et depuis le 1er janvier 2018 les Contrats Uniques d'Insertion.

Néanmoins, bien que des contrats aidés permettent de créer des emplois dans le secteur non-marchand, ceux-ci ont très peu d'impact sur l'emploi dans le secteur marchand, de plus les entreprises bénéficient d'un effet d'aubaine bien plus important dans le secteur marchand. En effet une étude a montré que, en 2017, dans le secteur non-marchand 1 contrat sur 5 (contre 6 contrats sur 10 dans le secteur marchand) bénéficiaient d'un effet d'aubaine ce qui conduit à une forme de dispersion d'argent public²⁰.

On peut aussi citer les lois Aubry qui exonèrent les employeurs d'une partie des charges sur les salaires jusqu'à 1,7 SMIC, ou le dispositif zéro charges de 2009, successeur des allègements Fillon de 2003 qui exonère jusqu'à 28 % des cotisations à travers une aide à l'embauche, destinée aux entreprises de moins de dix salariés pour les embauches pouvant bénéficier de la réduction générale de cotisations sur les bas salaires concernant les salariés percevant une rémunération inférieure à 1,6 fois le montant du SMIC.

18 <https://www.insee.fr/fr/statistiques/1377948?sommaire=1377956> consulté le 21/04/2019

19 https://solidarites-sante.gouv.fr/IMG/pdf/rapport_d_evaluation_prime_d_activite.pdf consulté le 21/04/2019

20 <https://www.ccomptes.fr/sites/default/.../05-mise-en-oeuvre-contrats-aides-Tome-2.pdf> consulté le 21/04/2019

II. 5. - La réduction du coût du travail (principalement sur les bas salaires)

Le coût du travail est constitué des charges patronales qui sont les charges payées par les employeurs dans le cadre de la rémunération de leurs salariés, elles permettent notamment de financer la sécurité sociale (assurance vieillesse, chômage et maladie, allocations familiales...).

En France, aujourd'hui, les charges patronales représentent entre 25 et 42% du salaire brut d'un employé ce qui est relativement élevé si on les compare aux autres pays du monde.

Ce coût du travail élevé est vivement critiqué par de nombreux économistes et hommes politiques puisqu'il est l'une des principales cause de la faible compétitivité des entreprises françaises qui dégrade la balance commerciale qui est déficitaire depuis 2004. En 2018, les échanges commerciaux de marchandises de la France enregistrent un déficit de 59,9 milliards d'euros.

C'est pourquoi l'Etat français cherche à baisser le coût du travail afin de rendre les entreprises plus compétitives et à les encourager à créer des emplois avec par exemple la mise en place du Crédit d'Impôt sur la Compétitivité et l'Emploi en 2013 qui permettait à toutes les entreprises employant des salariés à moins de 2,5% SMIC de bénéficier d'un crédit d'impôt égal à 6% des salaires correspondants.

Cependant les études faites depuis son établissement sur les effets de cette mesure montre que le CICE n'eut quasiment aucun impact sur le chômage ou l'investissement alors qu'elle coûte cher à l'Etat, près de 19 milliards d'euros en 2015.

Conclusion

Pour conclure, depuis les années 1980, le gouvernement français a radicalement modifié ses objectifs et la portée de ces politiques de l'emploi afin de s'adapter à la Mondialisation, aux mutations du marché du travail, et à la structure de l'économie française.

Néanmoins, bien que les politiques mises en place eurent un impact non négligeable sur l'emploi en France, la hausse du chômage observée et son maintien encore longtemps après la crise de 2008, révèle une forme d'inefficacité de ces politiques dans la lutte contre le chômage, les meilleurs résultats qu'enregistrent d'autres pays en matière d'emploi comme l'Allemagne, confirment bien la défaillance de l'adaptation du marché du travail français aux mutations de l'économie.

Cela peut être expliqué par la complexité de la mise en œuvre de telle politiques en France, du fait de l'alternance des gouvernements de gauche et de droite qui affecte la cohérence des réformes qui sont appliquées les unes après les autres réduisant l'efficacité de l'ensemble des politiques instaurées par le différents gouvernements. Mais aussi par l'opposition des syndicats en France face aux réformes qui sont jugées « défavorables aux employés et aux chômeurs » est particulièrement virulente ce qui affaiblit l'homogénéité et l'efficacité des politiques de flexibilisation du travail.

Toutes ces politiques misent en place visent à modifier la structure du marché du travail, or depuis les années 1970 la conjoncture économique de la France est globalement mauvaise du fait d'une succession de crises économiques que le pays met du temps à résorber²¹.

C'est pourquoi il est primordial de connaître les variables qui influencent réellement le chômage à la hausse ou à la baisse. Aujourd'hui l'idée que le coût du travail est la principale cause du chômage en France semble dominer la sphère politique, or comme nous l'avons montré les politiques d'exonération des charges patronales et de flexibilisation du marché du travail ont eu un impact positif mais modéré sur la baisse du chômage. D'un côté les allègements de cotisations sociales patronales, surtout sur les bas salaires, sont efficaces contre la hausse du chômage, ils auraient créé 170 000 emplois de 1997 à 2002, mais d'un autre côté cela fait 10 ans que le chômage en France semble se stabiliser autour des 9-10% de la population active²².

Nous avons donc décider de nous intéresser à d'autres variables corrélées au taux de chômage, autres que celles sur lesquelles sont ciblées les politiques publiques des 30 dernières années. Nous allons donc nous intéresser ici à 4 variables; le taux de croissance économique, le taux de croissance de l'investissement, le taux de croissance de l'inflation, et le taux de croissance du le revenu disponible net des ménages. La détermination de la relation entre ces variables et le taux de chômage devrait nous permettre de comprendre le problème du chômage élevé en France sous un angle nouveau et d'imaginer des actions publiques plus efficaces, mieux ciblées et de mieux comprendre ce qui motive l'offre et la demande de travail.

21 Blanchon.B, 2015, *Histoire des faits économiques*, Chapitre 36 : Les chocs pétroliers de 1973 à 1979 p114, Chapitre 48 De la crise financière des subprimes à la crise de l'euro p150, consulté le 22/04/2019

22 Barou.V, Buridant.J, Figliuzzi.A, Montoussé.M, 2016, *Histoire des faits économiques*, Partie 4: 1970-2015 Le temps des ruptures p278

II. 6. - Justification théorique des variables explicatives

Le taux de croissance économique

Selon l'économiste François Perroux la croissance économique correspond à l'augmentation soutenue sur une ou plusieurs périodes d'un indicateur de dimension, le plus souvent utilisé est le PIB qui mesure la richesse créée chaque année dans un pays à travers la production de biens et services.

La première explication que l'on peut avancer pour expliquer l'augmentation du chômage depuis les chocs pétroliers, est la dégradation significative de la conjoncture économique. Le rythme de croissance faible que connaît la France depuis les années 1970 est une des explications de la baisse du nombre de création d'emplois et de la hausse du nombre de perte d'emplois sur le marché du travail²³.

Aujourd'hui dans la majeure partie des pays, ce sont les entreprises privées qui créent la grande majorité des emplois, car pour produire des biens et des services, elles ont besoin de deux facteurs: le facteur travail et le facteur capital, de ce fait, les aléas de la variation du taux de chômage vont dépendre directement des besoins en facteur travail des entreprises nationales.

Selon la théorie microéconomique, ces besoins vont évoluer en fonction du niveau de production des entreprises permettant de maximiser leurs profits, ainsi plus les entreprises souhaitent produire des quantités élevées, plus elles auront besoin de facteur travail, et plus elles créeront des emplois. Le rythme de création et de destruction d'emplois est alors fortement corrélée à "l'état de santé" des entreprises nationales et à leur niveau de production, ce qui traduit la conjoncture économique du pays. Afin de mieux analyser la relation entre la conjoncture économique et le chômage, l'économiste américain Arthur Okun a mis en avant une loi dans les années 1960, "la loi d'Okun", expliquant que le taux de chômage est une fonction croissante de l'écart entre le taux de croissance effectif et le taux de croissance potentiel, défini comme le taux de croissance compatible avec une absence de tensions sur le marché des biens et services et sur le marché du travail qui n'entraîne pas d'inflation tout en utilisant de manière optimal les facteurs de production disponibles²⁴.

23 Barou.V, Buridant.J, Figliuzzi.A, Montoussé.M,2016, *Histoire des faits économiques*, Partie 4: 1970-2015 Le temps des ruptures p205-206

24 [Okun.A](#), *Potential GNP: Its measurement and significance*, American Statistical Association, *Proceedings of the Business and Economics Section*, 1962, pp 98-103.

Lorsque le PIB effectif est inférieur au PIB potentiel cela traduit une utilisation des facteurs de production qui ne permet pas de maximiser la production nationale ce qui a pour conséquence un taux de chômage supérieur au taux de chômage naturel pouvant être atteint, qui correspond au taux de chômage en l'absence de tensions sur le marché du travail et qui ne génère pas d'inflation.

Plus généralement, la loi d'Okun montre l'existence d'une relation décroissante entre la variation du taux de croissance du PIB et la variation du taux de chômage.

Tout choc économique influençant le processus de production des entreprises influence indirectement le niveau de chômage du pays.

La loi d'Okun peut être vérifiée en France empiriquement, en effet durant la période des Trente Glorieuses (1950-1970) le taux de croissance du PIB était en moyenne de 5% avec un taux de chômage équivalent à son niveau de plein emploi, soit 2%. Or depuis les années 1970 la France connaît un taux de croissance moyen d'à peine 2% jusqu'à aujourd'hui qui c'est accompagné d'une augmentation rapide du chômage avoisinant les 9% en moyenne²⁵.

Le taux de croissance du PIB est donc une variable qui influence énormément le taux de croissance du chômage en France, c'est pourquoi il est pertinent de l'inclure dans notre modèle.

Le taux de variation de la FBCF

L'investissement est une opération qui consiste à acquérir certains biens de production (investissement matériel) ou à effectuer des dépenses (investissement immatériel) dans le but de développer le potentiel de production de l'entreprise dans le futur. Il existe trois types d'investissement: l'investissement de capacité (qui accroît le stock de capital), l'investissement de productivité (qui permet de moderniser et d'améliorer la productivité des facteurs de production), et l'investissement de remplacement (qui renouvelle le capital initial)²⁶.

L'investissement total des entreprises privées et publiques est mesurable par la FBCF (Formation Brute de Capital Fixe) qui correspond à l'acquisition moins la cession d'actifs fixes par les producteurs résidents. L'État et les ménages peuvent également investir, dans le cas des ménages, il s'agit uniquement des achats de logements.

25 <https://data.oecd.org/fr/gdp/previsions-du-pib-nominal.htm#indicator-chart> consulté le 22/04/2019

26 Bernard et Colli, *Dictionnaire économique et financier*, , sixième édition, Seuil.consulté le 22/04/2019

L'investissement des entreprises est considéré comme une clé de la croissance, car il rend plus efficace le travail humain, concrètement il se traduit par l'achat de machines, la construction et l'aménagement de bureaux, des dépenses de recherches...

L'investissement est au cœur de la croissance. La FBCF, qui est un indicateur permettant de mesurer, au moins approximativement, l'investissement, est très souvent utilisée par les économistes pour évaluer la bonne santé d'une économie : une FBCF qui s'accroît rapidement laisse penser que la croissance va être rapide. Entre 1982 et 2017, la FBCF représente en moyenne 22% du PIB français, bien que cela représente à peine un quart du PIB, cette composante est l'une des plus déterminante de la croissance car l'investissement est beaucoup plus volatile que la consommation et donc influence beaucoup plus le taux de croissance du PIB et donc l'évolution de l'emploi²⁷.

Comment peut-on justifier le rôle important de l'investissement dans la détermination PIB et du niveau chômage?

Plusieurs théories économiques expliquent les bienfaits de l'investissement sur la croissance et l'emploi par le biais d'un double effet : il influence à la fois la demande globale et l'offre globale.

Du côté de la demande globale c'est l'économiste J-M Keynes qui explique ce mécanisme à travers le multiplicateur d'investissement. D'après le multiplicateur keynésien, une augmentation de l'investissement entraîne une variation amplifiée du revenu national. En effet l'investissement est une composante de la demande globale composée de la demande de biens de consommation et de biens de production. Une fois dépensée, cette augmentation de la demande globale va se traduire en une augmentation du revenu des ménages dont une partie va être consommée, dépendant de leur propension marginale à consommer, ce qui va alimenter de nouveau la demande globale. Un accroissement de l'investissement va donc engendrer des vagues successives de revenus et de dépenses formant un cercle vertueux durant une période qui dépendra du niveau d'investissement initial. En effet, plus l'accroissement de l'investissement est élevé, plus l'effet du multiplicateur va être important. Pour J-M Keynes l'existence du chômage s'explique par la faiblesse de la demande globale de ce fait l'augmentation de la consommation des ménages engendrée par investissement permet de stimuler la croissance et donc de stimuler la création d'emplois²⁸.

27 <https://donnees.banquemondiale.org/indicateur/NE.GD.LFTOT.ZS?end=2017&locations=FR&start=1982> consulté le 21/04/2019

28 Généreux.J, *Introduction à la politique économique*, 1999, Chapitre 7: La politique budgétaire, Partie B : La théorie du multiplicateur, consulté le 22/04/2019

L'investissement agit également sur l'offre :

L'investissement agit sur l'offre de trois manières : il permet l'augmentation du capital et donc des capacités de production, il favorise l'innovation, il accroît la productivité et la compétitivité.

1/ Tout d'abord l'investissement des entreprises permet d'accroître la quantité de capital de l'entreprise ce qui permet de produire davantage, cela contribue à stimuler la croissance extensive.

2/ L'investissement est la principale source de l'innovation. C'est Schumpeter qui met en avant cette relation, pour lui, l'investissement stimule l'innovation qui elle-même engendre de la croissance économique. Schumpeter considère que la croissance économique s'explique essentiellement par le progrès technique qui augmente la productivité globale des facteurs (travail et capital). Ce progrès technique dépend des découvertes (invention) et de leur application dans le système productif (innovation) qui sont faites par des « entrepreneurs-innovateurs ». Des vagues d'innovations détruisent les produits et les procédés traditionnels pour relancer la croissance jusqu'à leur épuisement. Ainsi l'investissement dans de nouveaux biens ou de nouveaux procédés de production permet de relancer l'activité économique et d'enclencher un nouveau cycle de croissance réduisant ainsi le chômage²⁹.

3/ Enfin, dans les modèles théoriques de la croissance endogène de Romer, Lucas et Barro, l'investissement joue un rôle clef dans la productivité, grâce notamment à son effet sur le progrès technique. L'investissement public et privé améliore la productivité du secteur privé lorsqu'il est dirigé vers le développement du capital humain et technologique. En effet l'investissement dans les sources de progrès technique produit des externalités positives qui permettent l'existence de rendements d'échelle croissants, conduisant une augmentation de la croissance économique et donc de l'emploi à long terme. On remarque en effet que les taux de croissance les plus élevés sont ceux des pays où la R&D est forte³⁰.

4/ L'investissement de productivité permet d'accroître la productivité des entreprises nationales et donc leur compétitivité vis-à-vis des entreprises étrangères engendrant une augmentation des exportations et de l'emploi.

29 Schumpeter.J, *Les cycles des affaires* (Business Cycles: a Theoretical, Historical and Statistical Analysis of the Capitalist Process), 1939 consulté le 22/04/2019

30 Paul Romer, *Endogenous Technological Change*, 1990, Journal of Political Economy consulté le 22/4/2019

En conclusion, l'investissement est une clef essentielle de la croissance : il joue à la fois sur l'offre et la demande qui sont les déterminant essentiel des variations du niveau d'emploi dans le pays.

La variation du taux d'inflation

L'inflation est définie comme la hausse globale et cumulative du niveau général des prix en un lieu géographique donné et mesuré par l'indice des prix à la consommation.

L'inflation peut avoir plusieurs sources :

Tout d'abord la demande globale, si l'accroissement de la demande globale est supérieur à celle de l'offre globale, les prix augmentent.

Puis les coûts de production, lorsque les coûts du facteur travail ou du capital augmentent les entreprises répercutent cette hausse sur les prix de vente afin de conserver leur marges bénéficiaires, c'est ce qui a été observé lors des chocs pétroliers.

Depuis le milieu des années 1980, la France connaît un faible taux d'inflation d'environ 2% qui depuis crise de 2008 a encore diminué jusqu'à frôler la déflation, alors que durant les Trente Glorieuses le pays était au plein emploi avec une inflation élevée d'environ 5%. On a pu observer empiriquement que la chute brutale du niveau général des prix après la crise des chocs pétroliers s'est accompagné d'une augmentation sans précédent du niveau de chômage en France et qui depuis cette période n'a jamais pu passer en dessous la barre des 7%. Mis à part la période des chocs pétroliers, on observe qu'il y a une relation négative entre l'inflation et le chômage en France³¹.

On peut donc se demander si l'inflation a une influence sur le niveau de chômage d'un pays. Pour expliquer cette relation Phillips c'est intéressé à l'existence d'une relation décroissante entre le chômage et la variation des salaires nominaux grâce à une étude empirique au Royaume-Uni de 1861 à 1957. Puis en 1960 Samuelson et Solow reprennent l'étude de Phillips et construisent une courbe appelée "la courbe de Phillips" et montre cette fois si une relation décroissante entre le taux d'inflation et l'écart entre le taux de chômage effectif et le taux de chômage naturel³².

31 <https://data.oecd.org/fr/price/inflation-ipc.htm> consulté le 22/04/2019

32 Phillips A. W. H. , *The Relation between Unemployment and the Rate of Change of Money Wage. Rates in the UK, 1861-1957*, *Economica*, 1958, consulté 22/04/2019

Taux de variation du le revenu disponible net

En économie, le revenu disponible est le revenu dont dispose effectivement un ménage afin de consommer ou d'épargner, il est composé de toutes les sources de revenu dont dispose les ménages comme les salaires, les revenus de transfert, les revenus non-salariaux (revenu du patrimoine), auquel on retire les prélèvements obligatoires. Le revenu disponible est la principale source de consommation des ménages or la demande des ménages est la principale composante de la demande globale et du PIB, c'est pour cette raison que le taux de croissance du revenu disponible net influence l'activité économique et donc indirectement le taux de chômage du pays.

Comme on a déjà évoqué précédemment, Phillips c'était déjà intéressé à la relation décroissante entre les salaires nominaux, qui est l'un des composants principaux du revenu disponible des ménages, et le taux de chômage, mais c'est J-M Keynes expliquera la pertinence d'accroître le revenu disponible des ménages afin de relancer l'activité économique. Pour l'économiste J-M Keynes, lorsque que l'économie est en récession, cela est principalement due à la faiblesse de la demande des ménages or celle-ci dépend surtout du revenu dont il dispose, de ce fait lorsque le taux de croissance du revenu disponible des ménages net ralentit, la demande globale diminue, ce qui pénalise la croissance et la création d'emplois. Pour relancer l'activité économique et réduire le chômage, il est nécessaire d'accroître le revenu disponible des ménages grâce notamment à une augmentation des revenus de transferts ou à une baisse des impôts. En effet une partie de cette hausse des revenus va être consommée par les ménages, ce qui va relancer la conjoncture économique³³.

En France, le taux de croissance du revenu disponible net des ménages est très instable depuis les chocs pétroliers du notamment à la succession des périodes de récession qu'à connu le pays jusqu'à la crise de 2008 qui a considérablement ralenti celui-ci tout en s'accompagnant d'une hausse du chômage.

Les théories économiques et les données statistiques sur la France montre bien une relation négative entre la variation du revenu disponible net des ménage et la vacation du taux de chômage.

33 Génereux.J, 2014, *Jacques Génereux explique l'économie à tout le monde*, Chapitre 7 : La logique keynésienne ou «l'économie de la demande» p97, consulté le 22/04/2019

III. - Le modèle théorique RLM

Le modèle de Régression Linéaire Multiple (RLM) est une équation économétrique destinée à évaluer la dépendance (et la force de cette dépendance) entre une variable à expliquer, notée Y et plusieurs variables explicatives, notées X_{it} . Ce modèle est très proche du modèle de Régression Linéaire Simple (RLS) qui met en relation une variable explicative avec une variable expliquée. Ces deux modèles fonctionnent de la même manière à la différence que la régression multiple a l'avantage d'expliquer Y en fonction de plusieurs variables explicatives. Ainsi les modèles de régression multiple seront d'une manière générale, plus fidèles à la réalité, l'évolution d'une variable ne dépendant pas d'une seule variable explicative mais souvent de plusieurs qui jouent toutes un rôle différent. Ces modèles portent l'appellation « linéaire » car la corrélation entre les variables sera représentée par une droite.

On note l'équation :

$$- Y_t = \alpha + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_n X_{nt} + \varepsilon_t \quad \forall t \in [1;N]$$

Avec :

- Y : la variable à expliquer
- t : date t
- α & β : paramètres à estimer grâce au modèle
- X_{it} : variable explicative i avec $i \in [1;k]$ (k étant le nombre de variables explicatives que nous avons décidé de prendre en compte, $k=1$ pour une RLS)
- N : le nombre d'observations
- ε : le terme d'erreur

Dans notre exemple Y_t représente le taux de chômage de l'année t , $t \in [1;36]$. Et nous tentons d'expliquer ce taux grâce à 4 variables explicatives que sont ; X_2 le taux de variation de la Formation Brute de Capital Fixe (FCBF), X_3 le taux de variation de l'inflation, X_4 le taux de variation du revenu disponible des ménages et X_5 le te taux de croissance du PIB.

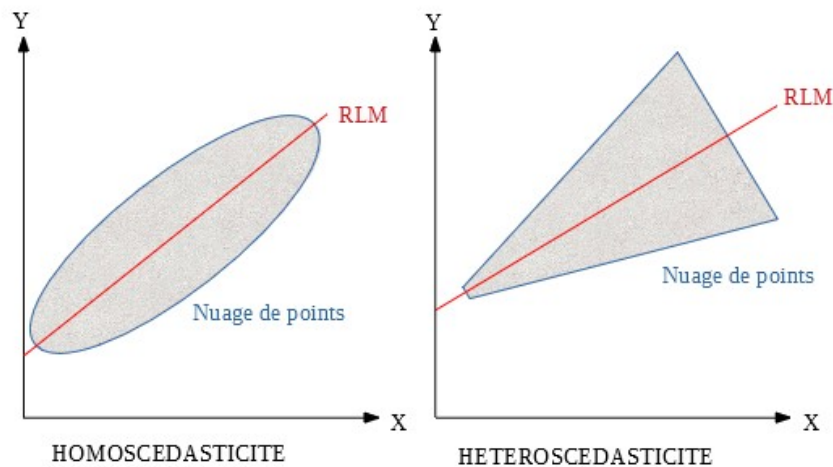
III. 1. – Les hypothèses du modèle RLM

Pour que la corrélation entre les variables soit avérée et puisse être interprétée, le modèle est soumis à 6 hypothèses différentes que voici :

- H_1 : $Y_t = \alpha + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_n X_{nt} + \varepsilon_t \quad \forall t \in [1;N]$. Le modèle est un modèle linéaire qui relie, les variables explicatives X à la variable expliquée Y .

- H_2 : $E[\varepsilon] = 0$. L'espérance du terme d'erreur est égale à 0. En effet le terme d'erreur représente la différence entre le \hat{Y} estimé grâce au modèle et le Y observé dans la réalité. Le modèle n'étant pas parfait $\hat{Y} \neq Y$ dans la plupart des cas. Cette différence peut être > 0 ou < 0 , mais la somme de toutes ces différences est égale à 0, car les différences positives et négatives s'annulent entre elles. Si cette hypothèse n'est pas vérifiée le modèle est inutilisable.

- H_3 : $\text{Var}[\varepsilon_t] = \sigma^2 \quad \forall t \in [1;N]$. il s'agit de l'**hypothèse d'homoscédasticité**, sous cette hypothèse ; la dispersion des résidus est homogène. C'est à dire que la distribution des résidus garde la même variance quand X augmente.



- H_4 : $\text{Cov}[\varepsilon_t, \varepsilon'_t] = 0$, il s'agit de l'**hypothèse de non-corrélation des termes d'erreur**. Autrement dit, les termes d'erreur ne sont pas corrélés entre eux. La valeur de ε_t n'a aucune influence sur la valeur ε'_t . Ce qui nous permet de dire que les Y sont indépendants entre eux.

- H_5 : C'est l'**hypothèse de non-multicolinéarité**. Elle signifie qu'aucune variable explicative n'est parfaitement corrélée à une autre. Bien sur, si cette hypothèse n'est pas respectée, il serait

impossible de déterminer l'impact individuel de chaque variable sur la variable expliquée. Or c'est précisément ce qui nous intéresse ici. Si deux variables explicatives sont parfaitement corrélées elles vont évoluer de manière strictement identique et il sera impossible de savoir laquelle influe sur Y.

III. 2. - Caractéristiques des estimateurs ponctuels.

Nous allons désormais nous intéresser aux caractéristiques des estimateurs ponctuels **a** et **b**. Nous verrons comment ils sont construits dans un premier temps, et pourquoi ces deux estimateurs sont les meilleurs que l'on puisse obtenir. Pour des questions d'intelligibilité nous allons procéder par étapes. En premier lieu nous ferons ces démonstrations sur la base d'une RLS, puis nous réitérerons cette manœuvre avec un modèle de RLM. Les démonstrations avec l'un ou l'autre modèle se ressemblent beaucoup, mais la RLM nous oblige à passer par des matrices et la formalisation est moins abordable. En revanche la méthode est identique.

III. 2. 1 - Construction des estimateurs

La méthode consiste à minimiser la somme des résidus (notée S).

$$S = \sum \varepsilon_i = \sum [y_i - bx_i - a]^2$$

$$(1) \frac{\partial S}{\partial a} = 0$$

$$(2) \frac{\partial S}{\partial b} = 0$$

$$(1) \sum x_i y_i - b \sum x_i^2 - a \sum x_i = 0$$

$$(2) \bar{y} = b \bar{x} - a = 0$$

$$(1) \hat{b} = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

$$(2) \hat{a} = \bar{y} - \hat{b} \bar{x}$$

Nous avons construit nos deux estimateurs a et b, mais pour pouvoir les interpréter, nous avons besoin qu'ils soient sans biais et efficaces. Sans biais signifie que l'on a : $E[\hat{b}] = \beta$.

Efficace signifie que la variance de l'estimateur est la plus petite possible. Si ces deux conditions sont réunies, les estimateurs seront les **BLUE's** (Best Linear Unbiased Estimators), les meilleurs estimateurs possibles. Nous pourrons ainsi les utiliser dans des intervalles de confiance.

III. 2. 2 - Propriétés des estimateurs

Démonstration : b est sans biais :

$$\text{On sait que: } y_i - \bar{y} = \beta (x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon})$$

$$\hat{b} = \frac{\sum (x_i - \bar{x}) [\beta (x_i - \bar{x}) + (\varepsilon_i - \bar{\varepsilon})]}{\sum (x_i - \bar{x})^2}$$

$$\frac{\beta \sum (x_i - \bar{x})^2 + \sum (x_i - \bar{x}) (\varepsilon_i - \bar{\varepsilon})}{\sum (x_i - \bar{x})^2}$$

$$\beta + \frac{\sum (x_i - \bar{x}) (\varepsilon_i - \bar{\varepsilon})}{\sum (x_i - \bar{x})^2}$$

On prouve facilement que $\bar{\varepsilon} \sum (x_i - \bar{x}) = 0$, grâce à l'hypothèse $E[\varepsilon_i] = 0$ nous obtenons ainsi :

$$\hat{b} = \beta + \frac{\sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2}$$

$$E[\hat{b}] = E[\beta] + E\left[\frac{\sum (x_i - \bar{x}) \varepsilon_i}{\sum (x_i - \bar{x})^2}\right]$$

$$\text{On appelle } \frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2}, w_i$$

$$E[\hat{b}] = \beta + \sum w_i \times E[\varepsilon_i], \text{ on sait que } E[\varepsilon_i] = 0, \text{ donc}$$

$$E[\hat{b}] = \beta$$

Nous avons prouvé que b est sans biais, son espérance est égale à β . Nous allons désormais démontrer que b est efficace, nous voulons prouver que $\lim_{n \rightarrow +\infty} \text{Var}[\hat{b}] = 0$

$$\hat{b} = \beta + \sum w_i \varepsilon_i$$

$$\text{Var}[\hat{b}] = E[(\sum w_i \varepsilon_i)^2]$$

$$\text{Var}[\hat{b}] = E[\sum w_i^2 \varepsilon_i^2 + 2 \sum w_i w'_i \varepsilon_i \varepsilon'_i]$$

$$\text{Var}[\hat{b}] = \sum w_i^2 E[\varepsilon_i^2] + 2 \sum w_i w'_i E[\varepsilon_i \varepsilon'_i]$$

Mais, par hypothèse : $E[\varepsilon_i^2] = \text{Var}[\varepsilon_i] = \sigma^2$ (homoscédasticité)
et $\text{Cov}[\varepsilon_i \varepsilon'_i] = E[\varepsilon_i \varepsilon'_i] = 0$ (non corrélation des termes d'erreur)

$$\text{Var}[\hat{b}] = \sigma^2 \sum w_i^2$$

$$\sum w_i^2 = \sum \left[\frac{(x_i - \bar{x})}{\sum (x_j - \bar{x})} \right]^2$$

$$\sum w_i^2 = \frac{1}{(\sum (x_j - \bar{x})^2)^2} \sum (x_i - \bar{x})^2$$

$$\text{Var}[\hat{b}] = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$$

σ^2 est un scalaire. Cependant $\lim_{n \rightarrow \infty} \sum (x_i - \bar{x})^2 = \infty \rightarrow \lim_{n \rightarrow \infty} \text{Var}[\hat{b}] = 0$. Nous prouvons ainsi que b est un estimateur efficace et sans biais de β .

Dans le cadre d'une RLM la démarche à suivre est la même, minimiser la somme des écarts en se servant de la condition d'ordre 1. Puis ensuite il convient de montrer que l'estimateur obtenu

par MCO est bien le meilleur possible. En revanche en RLM tout sera présenté sous forme de matrices ce qui complexifie nettement l'écriture ? Quelques étapes de calcul sont ainsi survolées.

Forme matricielle d'une équation de RLM :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & \dots & x_{1,1} & \dots & x_{k,1} & \dots \\ 1 & \dots & x_{1,2} & \dots & x_{k,2} & \dots \\ \dots & & \dots & & \dots & \\ 1 & \dots & x_{1,n} & \dots & x_{k,n} & \dots \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{bmatrix}, \mu = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix}$$

b doit alors vérifier ;

$$b \in \underset{\beta \in \mathbb{R}^{k+1}}{\operatorname{argmin}} \|Y - X\beta\|^2$$

b est alors unique et est donnée par : $b = (M' M)^{-1} M' Y$

$$Y = M\beta + (\mu - \bar{\mu})$$

$$b = (M' M)^{-1} [M\beta + (\mu - \bar{\mu})] \rightarrow b = \beta + (M' M)^{-1} M' (\mu - \bar{\mu})$$

$$M' \bar{\mu} = \begin{bmatrix} \dots & x_{2,\dots} & \dots \\ \dots & \dots & \dots \\ \dots & x_{k,\dots} & \dots \end{bmatrix} \begin{bmatrix} \bar{\varepsilon} \\ \dots \\ \bar{\varepsilon} \end{bmatrix} = 0$$

Comparons avec la RLS :

$$\text{RLM : } b = \beta + (M' M)^{-1} M' \mu \quad \text{RLS : } \hat{b} = \beta + \sum w_i \varepsilon_i$$

Vérifions que b est sans biais et efficace :

$$E[b] = \beta + (M' M)^{-1} M' E[\mu] \quad \text{Or, } E[\mu] = 0$$

$$E[b] = \beta \quad b \text{ est bel et bien sans biais}$$

Est-il efficace ?

$$\text{Cov}[b] = E[(b - \beta)(b - \beta)']$$

$$\text{Cov}[b] = \begin{bmatrix} b_2 - \beta_2 \\ \dots \\ b_k - \beta_k \end{bmatrix} [(b_2 - \beta_2) \dots (b_k - \beta_k)]$$

$$\text{Cov}[b] = E[(M' M)^{-1} M' \mu][\mu' M (M' M)^{-1}]$$

$$\text{Cov}[b] = \sigma^2 (M' M)^{-1}$$

Les estimateurs MCO sont les BLUE dans un modèle de RLM. A partir de ces résultats, nous allons pouvoir utiliser l'induction statistique (communément appelée extrapolation). Nous allons ainsi pouvoir tirer des conclusions générales sur le comportement des variables, conclusions qui sortent du cadre posé par les valeurs observées de l'échantillon. \hat{b} (resp \hat{a}) est un estimateur qui fonctionne pour cet échantillon en particulier mais quid de l'équation théorique ? Passer par un intervalle de confiance va nous permettre, en nous basant sur la valeur de \hat{b} (resp. \hat{a}), de déterminer un coefficient qui fonctionnera pour l'ensemble de la population mère. De plus c'est à partir de ces valeurs obtenues par intervalle de confiance que nous allons pouvoir effectuer différents tests statistiques qui nous en diront plus sur l'éventuelle significativité de chaque variable explicative.

III. 3. - Les estimateurs par Intervalle de Confiance (IC)

Avant toute chose, nous aurons besoin d'une hypothèse supplémentaire qui s'ajoute aux 5 premières qui sont discutées plus haut. L'hypothèse H_6 dite de **normalité des résidus** dit que les résidus du modèle sont distribués normalement, c'est à dire en suivant une loi Normale. On l'écrit formellement de la manière suivante :

$$\varepsilon_i \sim N(0; \sigma)$$

Cette hypothèse est primordiale car on sait que toute variable aléatoire étant une combinaison de variables aléatoires qui suivent une loi Normale, suivra aussi une loi Normale. Etant donné que Y dépend de ε , qui est la seule variable aléatoire du modèle, Y suit nécessairement une loi Normale. Et cette information va permettre l'induction statistique.

$$\mu \sim N(0; \sigma^2 I) \quad \text{puisque } \mu \text{ est le vecteur qui contient les valeurs de } \varepsilon \text{ distribuées normalement.}$$

$$\text{Cov}[\mu] = \sigma^2 I$$

$$b = \beta + (M' M)^{-1} M' \mu$$

$$\text{Sachant cela : } b \sim N(\beta, \sigma^2 (M' M)^{-1})$$

Si l'on centre et réduit on obtient : $\frac{b_j - \beta_j}{\sigma \sqrt{\gamma_{jj}}} \sim N(0,1)$ ou γ_{jj} représente un élément de la diagonale de $(M' M)^{-1}$

$$\text{et : } \frac{1}{\sigma^2} \varepsilon' \varepsilon \sim X^2(N-k)$$

Pour poursuivre notre démarche nous allons devoir rappeler que si Y suit une loi de X^2 à n degrés de liberté et que X suit une loi Normale centrée réduite alors :

$$T = \frac{X}{\sqrt{Y/n}} \sim \text{Student } T(n) \quad \text{avec } n \text{ le degré de liberté de la loi de Student récemment créée.}$$

Ainsi, on sait que l'on peut écrire :

$$\frac{\left[\frac{b_j - \beta_j}{\hat{\sigma} \sqrt{y_{jj}}} \right]}{\left[\frac{\sqrt{\varepsilon' \varepsilon / \sigma}}{\sqrt{N-k}} \right]} \sim t(N-k)$$

$$\text{Ce qui revient à écrire : } \left[\frac{b_j - \beta_j}{\hat{\sigma} \sqrt{y_{jj}}} \right] \sim t(N-k) \quad \forall j \in [2, k] \quad \text{car} \quad \hat{\sigma}^2 = \frac{\varepsilon' \varepsilon}{N-k}$$

$$\text{ou encore : } \left[\frac{b_j - \beta_j}{\hat{\sigma}_{bj}} \right] \sim t(N-k) \quad \text{car} \quad \hat{\sigma}_{bj}^2 = \hat{\sigma}^2 y_{jj}$$

Sachant cela, on peut calculer un IC en résolvant :

$$P[-t_{s/2} \hat{\sigma}_{bj} + b_j \leq \beta_j \leq +t_{s/2} \hat{\sigma}_{bj} + b_j] = 1 - s \quad s \text{ étant le seuil de risque choisi. } T_{s/2} \text{ se trouve dans la table de student.}$$

Cet encadrement de β va nous permettre d'effectuer le T-test de Student. Grâce à l'induction statistique nous allons pouvoir déterminer la significativité des variables du modèle. Le T-test s'articule comme suit :

$$H_0: \beta_j = 0$$

$$\rightarrow \text{Acceptée si } T_j = |b_j / \hat{\sigma}_{bj}| \leq t_{(s/2)}$$

$$\rightarrow \text{Refusée si } T_j \geq t_{(s/2)}$$

Cette hypothèse pose $\beta_j = 0$, autrement dit que la valeur du coefficient β_j est nulle. Cela signifie que la variable associée au coefficient β_j n'est pas significative ou qu'elle n'explique pas les variations de Y. Si cette hypothèse est rejetée, cela voudra dire que la variable est bien significative et qu'il y a une dépendance entre cette variable et Y. Si cette hypothèse est acceptée,

cela signifie que T_j se trouve à l'intérieur d'une distribution de Student « classique » qui représente l'indépendance des deux variables. Pour pouvoir dire que la variable est significative il faut que T_j se trouve à l'extérieur de cette distribution dans les zones où la probabilité devient très faible, soit au dessus de $t_{(s/2)}$. Ce $t_{(s/2)}$ est souvent fixé à 5 % mais cette valeur est arbitraire tout dépend de la précision et de la confiance que l'on désire.

Nous n'avons pas parlé du F-test, le test de Fisher-Snedecor car après l'avoir effectué, on est obligé de le compléter avec un T-test, on peut donc appliquer directement le T-test sans se soucier des résultats obtenus avec le F-test. Le F-test est basé sur le même principe ; extrapolation par intervalle de confiance puis on vérifie si 'lon se trouve dans la zone de rejet ou d'acceptation de l'hypothèse nulle. Ce test est applicable car Soit X suit une loi du khi-deux à n degrés de liberté et Y suit une loi de khi-deux à p degré de liberté, alors $F = [(X/n)/(Y/p)]$ suit une loi de Fisher de paramètre n et p noté $F(n, p)$. Les différences majeures entre les deux test résident dans le fait que F-test teste l'hypothèse dite jointe : $H_0 : \beta_1 = \beta_2 = \dots = \beta_n = 0$, et la distribution de Fisher n'est pas symétrique, le test est donc unilatéral. Ce test nous permet de dire, si H_0 est refusée, qu'au moins un des variables est significative. Mais impossible de savoir laquelle sans passer par un T-test par la suite. C'est pour cela que l'on peut appliquer directement le T-test individuel sur chaque variable pour en déterminer la significativité .

IV. - Analyse économétrique du sujet

IV. 1. - Définitions des variables

- **Y** : le taux de chômage
- **X2** : le taux de variation de la Formation Brute de Capital Fixe (FCBF) (en %)
- **X3** : le taux d'inflation annuel (en %)
- **X4** : le taux de variation du revenu disponible des ménages (en %)
- **X5** : le taux de variation du PIB (en %)

Toutes les données sont présentées de 1982 à 2017, toutes sont donc des séries temporelles.

Notre modèle initial est donc : $Y = \alpha - \beta_2 X_2 - \beta_3 X_3 - \beta_4 X_4 - \beta_5 X_5$

année	Y taux de chômage	X2 taux de variation FCBF	X3 taux de viation inflation	X4 taux de variations revenu disponible des ménages	X5 taux de croissance du PIB
1982	0,069	-0,009	0,118	0,0271	0,02509
1983	0,071	-0,03	0,096	0,0042	0,01256
1984	0,077	-0,008	0,074	-0,003	0,01524
1985	0,086	0,023	0,058	0,0137	0,01623
1986	0,086	0,042	0,027	0,0277	0,02351
1987	0,087	0,052	0,031	0,0173	0,0577
1988	0,084	0,087	0,027	0,0328	0,04734
1989	0,078	0,074	0,036	0,0346	0,04353
1990	0,076	0,046	0,034	0,0396	0,02914
1991	0,078	-0,005	0,033	0,0143	0,01039
1992	0,086	-0,015	0,023	0,0271	0,016
1993	0,096	-0,054	0,021	0,0097	-0,00613
1994	0,102	0,016	0,017	0,0066	0,02345
1995	0,096	0,013	0,019	0,0214	0,02085
1996	0,101	0,008	0,02	0,0119	0,01388
1997	0,103	0,008	0,012	0,0138	0,02337
1998	0,099	0,064	0,006	0,03	0,03556
1999	0,096	0,078	0,005	0,0283	0,03407
2000	0,081	0,066	0,017	0,0336	0,03875
2001	0,074	0,023	0,016	0,0305	0,01954
2002	0,075	-0,009	0,02	0,0307	0,0118
2003	0,081	0,019	0,021	0,0097	0,0082
2004	0,085	0,035	0,021	0,0208	0,02786
2005	0,085	0,029	0,019	0,0131	0,01608
2006	0,084	0,036	0,016	0,0257	0,02375
2007	0,077	0,055	0,015	0,0281	0,02361
2008	0,071	0,009	0,028	0,0047	0,00195
2009	0,087	-0,091	0,001	0,0174	-0,02941
2010	0,089	0,021	0,015	0,0174	0,01966
2011	0,088	0,021	0,021	0,0023	0,02079
2012	0,094	0,002	0,02	0,001	0,00183
2013	0,099	-0,008	0,009	-0,0061	0,00576
2014	0,103	0	0,005	0,0133	0,00948
2015	0,104	0,01	0	0,0105	0,01067
2016	0,101	0,028	0,002	0,0188	0,0188
2017	0,095	0,045	0,01	0,0146	0,01819

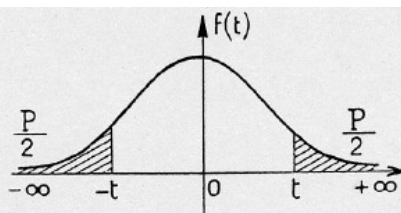
IV. 2. - Modèles intermédiaires

Le premier modèle que nous avons essayé contient toutes les variables. La RLM, obtenue avec Eviews donne les résultats suivants :

Dependent Variable: Y Method: Least Squares Date: 04/22/19 Time: 20:23 Sample: 1982 2016 Included observations: 35				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.096896	0.002993	32.37729	0.0000
X2	-0.135077	0.080011	-1.688224	0.1017
X3	-0.329188	0.059075	-5.572410	0.0000
X4	-0.383238	0.130046	-2.946933	0.0062
X5	0.419909	0.187314	2.241740	0.0325
R-squared	0.560083	Mean dependent var	0.087114	
Adjusted R-squared	0.501428	S.D. dependent var	0.010482	
S.E. of regression	0.007401	Akaike info criterion	-6.842789	
Sum squared resid	0.001643	Schwarz criterion	-6.620596	
Log likelihood	124.7488	Hannan-Quinn criter.	-6.766088	
F-statistic	9.548683	Durbin-Watson stat	0.944337	
Prob(F-statistic)	0.000042			

$$\hat{Y}_{(t)} = 0,096896 - 0,135077 X_{2(t)} - 0,329188 X_{3(t)} - 0,383238 X_{4(t)} + 0,419909 X_{5(t)}$$

Pour évaluer la significativité éventuelle des variables, nous aurons besoin de t-stat tabulé ($t_{(s/2)}$). Cette valeur se trouve dans la table de Student. Nous choisissons un risque α de 5 % et nous avons 35 observations. Nous pouvons en déduire que le degré de liberté à considérer pour le T-test est 30. En effet le degré de liberté du T-test est $(N-[k+1])$, k étant le nombre de variables explicatives et N le nombre d'observations.



$\frac{P}{2}$ v	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,388	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,611
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

En l'occurrence, nous allons comparer les t-stat de chaque variables avec la valeur de 2,042.

On constate alors que la variable X2 (taux de variation FCBF) n'est pas significative. La valeur de t-stat qui lui est associée est inférieure à la valeur t-stat tabulé ($2,042 > 1,68$). Ce n'est pas le seul problème ; R^2 est relativement faible ($= 0,501$), ce qui indique que la qualité du modèle est moyenne (les variables expliquent moins de 50 % des variations de Y). Et surtout le coefficient de X5 (0,42) est positif. En soit ce résultat n'est pas problématique mais son interprétation économique elle pose problème. Ce modèle indique que le taux de chômage évolue dans le même sens que le taux de variation du PIB. Plus le PIB augmente plus le taux de chômage augmente. Nous sommes à l'opposé de n'importe quelle théorie économique du chômage (telle que la relation d'Okun par exemple), la logique voudrait que la relation entre les deux soit négative.

Pour déterminer la source de cette incohérence, nous avons regardé la matrice des corrélations.

Correlation				
	X2	X3	X4	X5
X2	1.000000	-0.172237	0.509740	0.854948
X3	-0.172237	1.000000	-0.076466	0.107792
X4	0.509740	-0.076466	1.000000	0.518119
X5	0.854948	0.107792	0.518119	1.000000

Nous pouvons directement remarquer la très forte corrélation entre X5 et X2. Elle n'est pas parfaite ($0,85 < 1$) mais cette relation risque fortement de poser un problème de multicolinéarité. C'est peut être la cause de la dépendance anormale entre Y et X5.

Nous décidons ainsi de tester deux modèles supplémentaires ; un où l'on exclu X2 et un où l'on exclu X5.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 20:25
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.098784	0.002857	34.57150	0.0000
X3	-0.280184	0.052965	-5.290034	0.0000
X4	-0.405756	0.133165	-3.047025	0.0047
X5	0.148476	0.098931	1.500798	0.1435

R-squared	0.518290	Mean dependent var	0.087114
Adjusted R-squared	0.471673	S.D. dependent var	0.010482
S.E. of regression	0.007619	Akaike info criterion	-6.809174
Sum squared resid	0.001799	Schwarz criterion	-6.631420
Log likelihood	123.1605	Hannan-Quinn criter.	-6.747814
F-statistic	11.11801	Durbin-Watson stat	0.768828
Prob(F-statistic)	0.000041		

$$\hat{Y}_{(t)} = 0.098784 - 0,280184 X 3_{(t)} - 0,405756 X 4_{(t)} + 0,148476 X 5_{(t)}$$

A noter que, en théorie, nous avons ôté une variable, le degré de liberté de la loi de Student va donc changer. On passe à $35 - (3+1) = 31$. En réalité, la loi de Student étant prévue pour des petits échantillons, à partir de 30 degrés de liberté, la valeur de t-stat tabulé va rester la même, c'est à dire 2,042.

Très peu de changements cependant, X5 n'est toujours pas significatif et son coefficient est contraire à l'intuition. R² ajusté a diminué, il ne vaut désormais plus que 0,47 et l'écart type σ de régression a augmenté. D'une manière générale nous n'avons pas résolu notre problème et la qualité du modèle s'est empirée. Nous allons donc tenter la deuxième option ; nous enlevons X5 du modèle et ajoutons X2. Même si nous pensons retirer du modèle une variable essentielle X5 (le taux de croissance du PIB), nous tenons à voir ce que cela pourrait donner.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 20:27
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.099477	0.002936	33.88012	0.0000
X3	-0.262929	0.054368	-4.836070	0.0000
X4	-0.329761	0.135887	-2.426735	0.0212
X2	0.018880	0.043635	0.432681	0.6682

R-squared	0.486391	Mean dependent var	0.087114
Adjusted R-squared	0.436687	S.D. dependent var	0.010482
S.E. of regression	0.007867	Akaike info criterion	-6.745056
Sum squared resid	0.001919	Schwarz criterion	-6.567302
Log likelihood	122.0385	Hannan-Quinn criter.	-6.683695
F-statistic	9.785750	Durbin-Watson stat	0.607308
Prob(F-statistic)	0.000107		

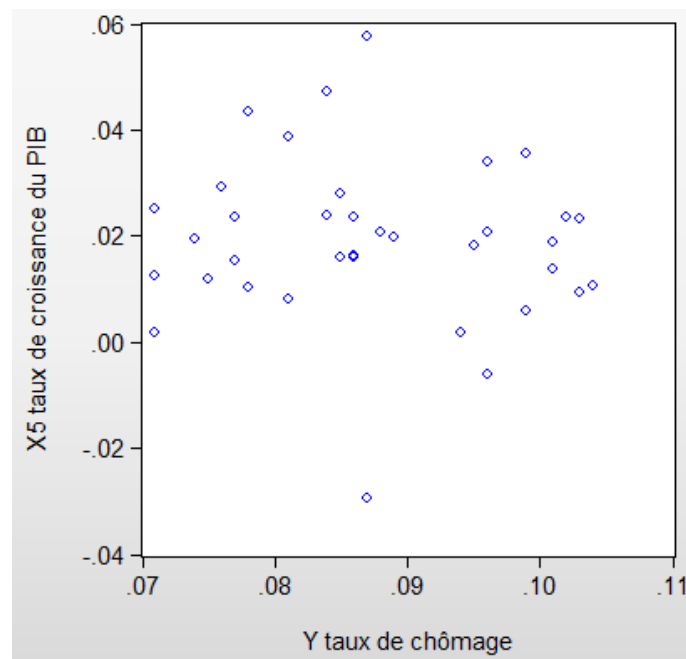
$$\hat{Y}_{(t)} = 0.099477 - 0.262929 X_{3(t)} - 0.329761 X_{4(t)} + 0.018880 X_{2(t)}$$

Force est de constater que ce modèle n'est pas non plus le bon. C'est désormais X2 qui d'une part n'est plus significatif et qui d'autre part évolue dans le même sens que le taux de chômage, alors que la théorie nous dit le contraire.

A ce moment, nous savons déjà qu'il va être difficile de trouver un modèle qui soit efficace. 2 des variables explicatives semblent très corrélées entre elles (X2 et X5) et cette corrélation entraîne des résultats aberrants vis à vis de la théorie économique. Il est envisageable que la théorie soit fausse et que notre modèle en soit la preuve. Mais cette hypothèse nous paraît irréaliste, pour essayer d'obtenir plus de précision dans le but d'améliorer notre modèle et de le rendre plus pertinents nous allons regarder de plus près la principale complication que nous avons observée : la relation incohérente entre X5 et Y.

En effet, lorsque nous avons réfléchi aux variables explicatives à intégrer pour expliquer le chômage en France, le taux de croissance du PIB nous est apparu comme un évidence. Arthur Okun avait déjà identifié (à l'aide d'une RLS) la corrélation négative très nette entre le taux de croissance du PIB et le taux de chômage. Cependant nous n'arrivons pas aux mêmes conclusions. C'est pourquoi à ce stade nous avons effectué quelques tests supplémentaires qui sortent quelque peu de l'analyse des résultats de la RLM pour comprendre d'où le problème pourrait venir.

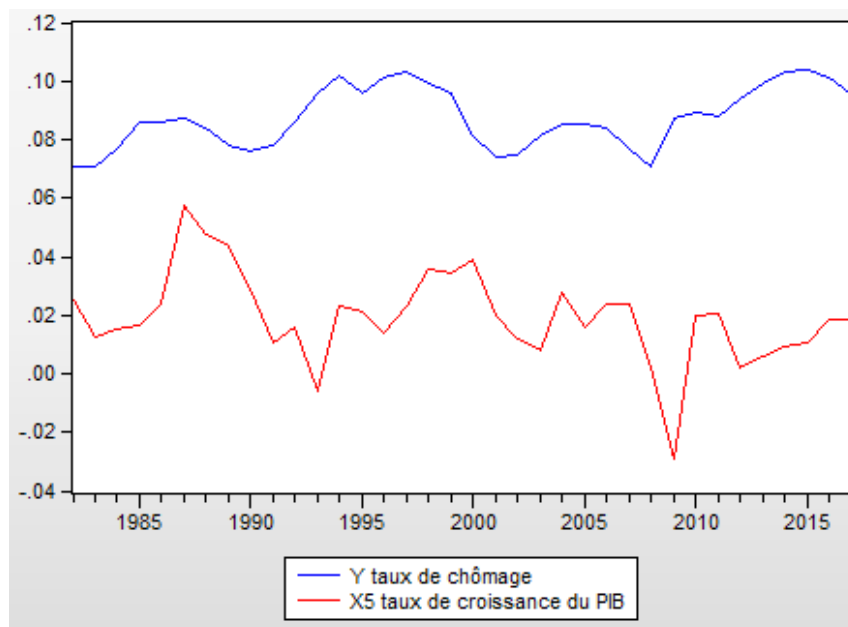
Pour commencer nous avons décidé d'observer la relation en RLS entre X5 et Y. A l'image d'Okun nous devrions trouver une corrélation négative très claire entre les deux.



On constate immédiatement que cette relation n'est pas aussi nette dans notre modèle. Ce, pour plusieurs raisons envisageables. La première serait que cette relation identifiée dans les années 60 ne fonctionne plus. Cependant cette théorie, qui consiste à dire que le taux de chômage est décroissant du taux de croissance du PIB est réputée pour être l'une des plus pérennes de la macroéconomie.

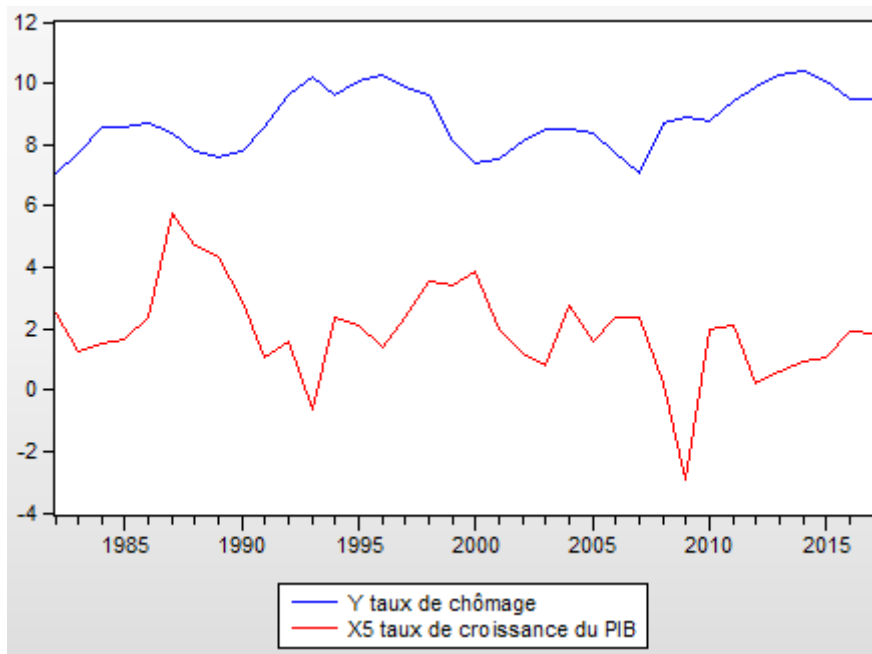
La seconde, plus probable viendrait de nos séries de variables, en effet en statistiques la qualité des données récoltées est primordiale et déterminante de la qualité finale des modèles. Nos données peuvent donc être erronées (hypothèse peu probable étant donné que les chiffres proviennent de sites gouvernementaux), ou alors la taille de notre échantillon n'est pas assez grande pour obtenir des résultats fiables. Il aurait été possible d'élargir notre échantillon à 40, 50 observations ou même plus, cependant certaines de nos variables ne sont pas disponibles sur d'aussi longues périodes. Notamment X4, le taux de variation du revenu disponible des ménages, qui n'est pas disponible pour des dates antérieures à 1980.

Nous allons donc devoir trouver une solution pour améliorer notre modèle sans élargir notre échantillon, et c'est en observant les graphiques que nous avons eu une première idée.



Nous avons alors constaté que la courbe de Y suit celle de X5 mais avec un léger décalage. En effet on peut deviner graphiquement une corrélation négative entre les deux mais il y a un décalage de quelques années entre les mouvements du chômage et ceux du PIB. Les variables macroéconomiques ne sont pas forcément très volatiles et les ajustements peuvent nécessiter du temps avant d'entraîner des résultats. Il serait peut être préférable, dans ce cas, d'expliquer les chiffres du chômage en $(t+1)$ avec les variables explicatives en (t) . C'est une manière artificielle de prendre en compte ce décalage temporel entre une augmentation du taux de variation du PIB et une baisse du taux de chômage ou réciproquement.

Graphiquement cela donne :



La relation paraît légèrement plus régulière, on peut vérifier que cette idée soit la bonne en relaçant les RLM effectuées précédemment avec ces nouvelles variables « décalées ».

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 17:22
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.099928	0.002656	37.62609	0.0000
X2	-0.173283	0.070993	-2.440838	0.0208
X3	-0.256324	0.052377	-4.893832	0.0000
X4	-0.457797	0.115361	-3.968393	0.0004
X5	0.308528	0.166179	1.856604	0.0732
R-squared	0.625138	Mean dependent var	0.087857	
Adjusted R-squared	0.575156	S.D. dependent var	0.010074	
S.E. of regression	0.006566	Akaike info criterion	-7.082252	
Sum squared resid	0.001293	Schwarz criterion	-6.860059	
Log likelihood	128.9394	Hannan-Quinn criter.	-7.005551	
F-statistic	12.50735	Durbin-Watson stat	1.209091	
Prob(F-statistic)	0.000004			

$$\hat{Y}_{(t+1)} = 0.099928 - 0.173283 X_{2(t)} - 0.256324 X_{3(t)} - 0.457797 X_{4(t)} + 0.308528 X_{5(t)}$$

Bien que le problème de X5 ne soit pas résolu (il est toujours corrélé à Y dans un sens contraire à l'intuition) R^2 est beaucoup plus intéressant désormais (0,575). En définitive nous avons réussi à faire augmenter R^2 , nous allons donc poursuivre dans cette direction et observer les résultats de différents modèles.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 17:26
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.102355	0.002652	38.59098	0.0000
X3	-0.193600	0.049154	-3.938660	0.0004
X4	-0.486793	0.123583	-3.939002	0.0004
X5	-0.039650	0.091812	-0.431856	0.6688
R-squared	0.550694	Mean dependent var	0.087857	
Adjusted R-squared	0.507213	S.D. dependent var	0.010074	
S.E. of regression	0.007072	Akaike info criterion	-6.958249	
Sum squared resid	0.001550	Schwarz criterion	-6.780495	
Log likelihood	125.7694	Hannan-Quinn criter.	-6.896888	
F-statistic	12.66510	Durbin-Watson stat	0.849206	
Prob(F-statistic)	0.000014			

$$\hat{Y}_{(t+1)} = 0.102355 - 0.193600 X_{3(t)} - 0.486793 X_{4(t)} - 0.039650 X_{5(t)}$$

On obtient des résultats intéressants. Le coefficient de X5 est désormais négatif, comme envisagé par la théorie économique, bien que son t-stat soit inférieur au t-stat tabulé, et que la variable soit par conséquent non-significative, nos résultats commencent à ressembler à ce que l'on espérait. Il nous reste alors deux pistes d'amélioration : utiliser la valeur du PIB en euros constants au lieu de son taux de variation, ou continuer à modifier la temporalité de nos séries.

Nous allons procéder dans cet ordre, d'abord en essayant de changer la nature de X5. Comme nous allons le constater cette solution s'avère très efficace et résout tous les problèmes de significativité et d'interprétation des modèles. En revanche cette solution semble très peu appropriée, en effet le PIB en euros constant suit une évolution quasiment linéaire, de ce fait il sera moins compliqué d'obtenir une corrélation entre X5 et Y. Mais étant donné que les variations de X5 sont quasiment uniformes (X5 est croissant sur toute la période) la corrélation est bien moins intéressante à analyser.

Pour commencer nous allons relancer notre modèle avec le PIB cette fois en euros constants et le taux de chômage de l'année (t), nous essayerons ensuite avec le taux de chômage de l'année (t+1).

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 17:31
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.127840	0.012257	10.43025	0.0000
X2	0.012515	0.040786	0.306855	0.7611
X3	-0.388380	0.073237	-5.303069	0.0000
X4	-0.411788	0.131365	-3.134695	0.0038
X5	-0.014013	0.005903	-2.374074	0.0242
R-squared	0.567624	Mean dependent var	0.087114	
Adjusted R-squared	0.509974	S.D. dependent var	0.010482	
S.E. of regression	0.007337	Akaike info criterion	-6.860078	
Sum squared resid	0.001615	Schwarz criterion	-6.637886	
Log likelihood	125.0514	Hannan-Quinn criter.	-6.783377	
F-statistic	9.846006	Durbin-Watson stat	0.806745	
Prob(F-statistic)	0.000033			

$$\hat{Y}_{(t)} = 0,127840 + 0,012515 X_{2(t)} - 0,388380 X_{3(t)} - 0,411788 X_{4(t)} - 0,014013 X_{5(t)}$$

Les résultats sont en effet très intéressants, X5 évolue dans le sens logique, R² est supérieur à 0,50, bien que de peu. Toutes les variables sont significatives sauf X2, vérifions si ce problème est dû à la multicolinéarité entre X2 et X5.

Correlation				
	X2	X3	X4	X5
X2	1.000000	-0.172237	0.509740	-0.017403
X3	-0.172237	1.000000	-0.076466	-0.691751
X4	0.509740	-0.076466	1.000000	-0.191553
X5	-0.017403	-0.691751	-0.191553	1.000000

Apparemment non, la forte corrélation entre X5 et X2 a disparu avec la transformation de X5. Cependant X2 est désormais corrélé positivement à X4, nous allons enlever la variable X2 et observer si la qualité de notre modèle s'en trouve améliorée.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 17:33
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.128004	0.012065	10.60971	0.0000
X3	-0.391609	0.071410	-5.483956	0.0000
X4	-0.392478	0.113615	-3.454468	0.0016
X5	-0.014132	0.005803	-2.435277	0.0208
R-squared	0.566267	Mean dependent var	0.087114	
Adjusted R-squared	0.524293	S.D. dependent var	0.010482	
S.E. of regression	0.007229	Akaike info criterion	-6.914087	
Sum squared resid	0.001620	Schwarz criterion	-6.736333	
Log likelihood	124.9965	Hannan-Quinn criter.	-6.852727	
F-statistic	13.49083	Durbin-Watson stat	0.787640	
Prob(F-statistic)	0.000008			

$$\hat{Y}_{(t)} = 0,128004 - 0,391609 X_{3(t)} - 0,392478 X_{4(t)} - 0,014132 X_{5(t)}$$

En utilisant le PIB en euros constant et en enlevant X2, on arrive à un modèle relativement fiable. Toutes les variables sont significatives (on rappelle que t-stat tabulé est 2,042), les coefficients sont conformes à l'intuition. Notre R^2 , d'une valeur de 0,52 est satisfaisant et notre σ de régression est assez faible. Nous allons donc conserver ce modèle à part, tester le même mais en utilisant les valeurs de $Y(t+1)$, voir si la qualité du modèle s'améliore.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 17:35
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.128567	0.010794	11.91095	0.0000
X3	-0.312754	0.063889	-4.895281	0.0000
X4	-0.600809	0.101648	-5.910661	0.0000
X5	-0.013009	0.005192	-2.505661	0.0177
R-squared	0.624117	Mean dependent var	0.087857	
Adjusted R-squared	0.587742	S.D. dependent var	0.010074	
S.E. of regression	0.006468	Akaike info criterion	-7.136676	
Sum squared resid	0.001297	Schwarz criterion	-6.958922	
Log likelihood	128.8918	Hannan-Quinn criter.	-7.075315	
F-statistic	17.15751	Durbin-Watson stat	1.333437	
Prob(F-statistic)	0.000001			

$$\hat{Y}_{(t+1)} = 0,128567 - 0,312754 X_{3(t)} - 0,600809 X_{4(t)} - 0,013009 X_{5(t)}$$

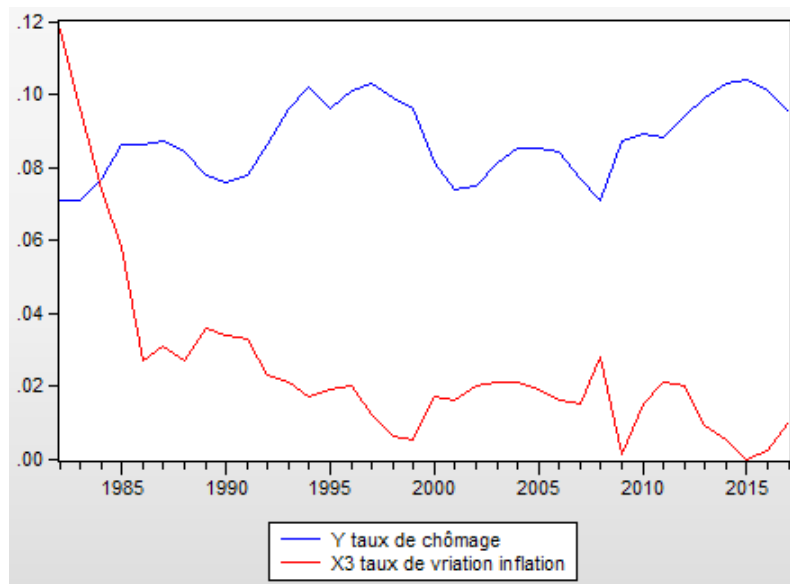
Nous obtenons un deuxième modèle encore plus satisfaisant que le premier. La significativité des variables a augmenté (les t-stats pour chaque variable sont plus grands que précédemment). R^2 est passé de 0,52 à 0,58, nos variables explicatives expliquent désormais quasiment 60 % des variations de Y. Et notre écart type a diminué, passant de 0,007 à 0,006.

A ce stade nous avons conclu que :

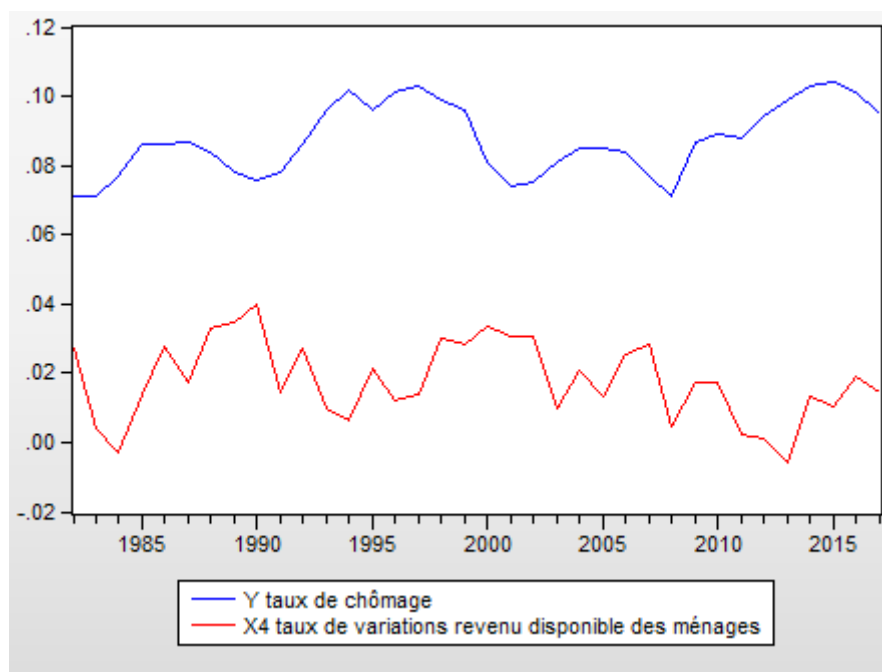
- X2 est à exclure du modèle, cette variable détériore la qualité de tous les modèles que nous avons testés jusqu'à présent.
- modifier la temporalité des séries de données semble au contraire avoir un impact positif sur la qualité de tous les modèles testés jusqu'à présent.

Nous nous rappelons que la relation entre le PIB en euros constants et le taux de chômage n'est pas la plus pertinente (car l'évolution du PIB en euros constants est quasiment linéaire) mais il est désormais clair que nous avons des modèles satisfaisants, si nous ne pouvons plus les améliorer, ils sont au moins utilisables.

Nous allons donc tester la dernière solution qui consiste à reprendre le taux de variation du PIB en guise de X5 et décaler certaines variables de 1 voire 2 années. Pour savoir lesquelles décaler et de combien de temps nous nous sommes appuyés sur les graphiques. Nous regardons si les évolutions de Y apparaissent en même temps ou plus tard que les variations des variables explicatives.



Il est difficile de déterminer graphiquement quelles variables auraient besoin d'être décalées dans le temps. En effet les variations de X4 et X3, sont assez irrégulières, leur lien avec Y est avéré, nous l'avons vu dans les régressions précédentes (leur significativité était démontrée par t-stat).



Pour ne pas détériorer la modèle nous avons décider de les laisser inchangées (c'est à dire de ne pas les décaler dans le temps).

En revanche, nous voulons absolument tester le modèle en décalant X5.

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 18:42
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.099944	0.002898	34.49012	0.0000
X3	-0.255417	0.053365	-4.786209	0.0000
X4	-0.219313	0.130508	-1.680464	0.1029
X5	-0.121463	0.096528	-1.258320	0.2177
R-squared	0.508399	Mean dependent var	0.087114	
Adjusted R-squared	0.460825	S.D. dependent var	0.010482	
S.E. of regression	0.007697	Akaike info criterion	-6.788849	
Sum squared resid	0.001836	Schwarz criterion	-6.611095	
Log likelihood	-122.8049	Hannan-Quinn criter.	-6.727489	
F-statistic	10.68642	Durbin-Watson stat	0.445504	
Prob(F-statistic)	0.000055			

$$\hat{Y}_{(t)} = 0,99944 - 0,255417 X 3_{(t)} - 0,219313 X 4_{(t)} - 0,121463 X 5_{(t+1)}$$

On se rapproche de la solution, X5 n'est pas significatif mais son coefficient est conforme à la logique.

A ce stade deux solutions s'offrent à nous ; nous fier à la théorie et choisir un modèle qui la confirme, ou nous fier à la statistique et réfuter la théorie qui dit que le taux de chômage doit diminuer si le taux de croissance du PIB augmente.

Pour nous aider dans notre prise de décision, nous avons décider de tester le modèle sur plusieurs sous-périodes et voici ce que nous avons trouvé, en utilisant $Y(t+1)$ et $X_i(t)$ avec 19 observations (1982-2000):

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 15:44
Sample: 1982 2000
Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.109597	0.003898	28.11949	0.0000
X3	-0.242816	0.048061	-5.052251	0.0001
X4	-0.330600	0.147682	-2.238602	0.0408
X5	-0.275767	0.113938	-2.420322	0.0287
R-squared	0.725226	Mean dependent var		0.087211
Adjusted R-squared	0.670271	S.D. dependent var		0.010326
S.E. of regression	0.005929	Akaike info criterion		-7.233184
Sum squared resid	0.000527	Schwarz criterion		-7.034355
Log likelihood	72.71525	Hannan-Quinn criter.		-7.199534
F-statistic	13.19676	Durbin-Watson stat		1.510175
Prob(F-statistic)	0.000175			

$$\hat{Y}_{(t+1)} = 0,109597 - 0,242816 X_{3(t)} - 0,330600 X_{4(t)} + 0,275767 X_{5(t)}$$

Bien que l'échantillon soit trop petit (19 observations) pour que le modèle puisse être retenu on remarque que sur la période 1982-2000 le modèle fonctionne parfaitement, c'est donc la période 2000-2017 qui pose problème. Ici nous utilisons $Y_{(t+1)}$ et $X_{i(t)}$ avec les 18 autres observations (2000-2017) ;

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 20:41
Sample: 2000 2017
Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.108035	0.002818	38.33265	0.0000
X3	-0.632360	0.150790	-4.193647	0.0009
X4	-0.748999	0.114053	-6.567101	0.0000
X5	0.074971	0.094021	0.797377	0.4385
R-squared	0.826635	Mean dependent var		0.088167
Adjusted R-squared	0.789486	S.D. dependent var		0.010188
S.E. of regression	0.004674	Akaike info criterion		-7.700297
Sum squared resid	0.000306	Schwarz criterion		-7.502436
Log likelihood	73.30267	Hannan-Quinn criter.		-7.673014
F-statistic	22.25156	Durbin-Watson stat		2.473409
Prob(F-statistic)	0.000014			

$$\hat{Y}_{(t+1)} = 0,108035 - 0,632360 X_{3(t)} - 0,748999 X_{4(t)} + 0,074971 X_{5(t)}$$

On constate tout de suite que dans cette configuration, R^2 ajusté passe à une valeur de 0,79, le t-stat de X4 (le taux de variation du revenu disponible des ménages) triple ; 2,24 sur la première période, 6,57 sur la seconde. Tandis que le t-stat de X5 a non seulement diminué (de 2,42 à 0,79 entre les deux périodes). Cela veut peut être dire que le revenu disponible des ménages est devenu une variable plus importante dans la détermination du taux de chômage. Et qu'à l'inverse le taux de croissance a perdu en significativité au cours du temps.

Si l'interprétation des informations supplémentaires que nous donnons est exacte, cela résout notre problème. A cause de la période 2000-2017 il est impossible de mettre en relation Y et X5 car la relation qui fonctionnait jusqu'ici s'arrête. Alors que sur toute la période 1982-2017 les deux autres variables majeures X3 (le taux d'inflation) et X4 (le taux de variation du revenu disponible des ménages) sont constamment significatives :

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 21:00
Sample: 1982 2016
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.102200	0.002594	39.39138	0.0000
X3	-0.197317	0.047775	-4.130104	0.0002
X4	-0.515172	0.103324	-4.985969	0.0000
R-squared	0.547991	Mean dependent var		0.087857
Adjusted R-squared	0.519740	S.D. dependent var		0.010074
S.E. of regression	0.006981	Akaike info criterion		-7.009394

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 21:02
Sample: 2000 2017
Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.107815	0.002771	38.91414	0.0000
X3	-0.591117	0.139912	-4.224935	0.0007
X4	-0.707274	0.100103	-7.065482	0.0000
R-squared	0.818762	Mean dependent var		0.088167
Adjusted R-squared	0.794597	S.D. dependent var		0.010188
S.E. of regression	0.004617	Akaike info criterion		-7.766994

Dependent Variable: Y
Method: Least Squares
Date: 04/22/19 Time: 21:01
Sample: 1982 2000
Included observations: 19

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.106521	0.004207	25.32092	0.0000
X3	-0.238198	0.054831	-4.344232	0.0005
X4	-0.528148	0.140525	-3.758389	0.0017
R-squared	0.617918	Mean dependent var		0.087211
Adjusted R-squared	0.570158	S.D. dependent var		0.010326
S.E. of regression	0.006770	Akaike info criterion		-7.008762

C'est particulièrement le cas sur la dernière période où le R^2 du modèle est de 0,79, le t-stat de X_4 est largement supérieur au t-stat tabulé ($7,07 > 2,042$). Quant à X_3 elle est elle aussi largement significative, tandis que l'écart type de régression n'est que de 0,005.

Ce modèle à 2 variables est de loin le meilleur que l'ont ai pu réaliser. La trop forte corrélation entre $X_{2(t)}$ et $X_{5(t)}$ nous empêche de les inclure dans le même modèle, et $X_{5(t)}$ et $Y_{(t+1)}$ ne sont pas linéairement corrélées. La variable $X_{5(t)}$ ne peut pas être considérée comme significative bien qu'elle le soit sur la première moitié des observations. $X_{2(t)}$ ne peut que réduire la qualité du modèle comme nous l'avons observé, il ne nous reste que 2 variables explicatives pour expliquer le chômage. Conscients que le principe de la RLM est d'omettre le moins de variables explicatives que possible dans notre modèle, aux vues des résultats obtenus lors de cette dernière régression, nous pensons judicieux de préférer le modèle à 2 variables. En ajouter une, que ce soit $X_5(t)$ ou X_2 donne des résultats moins flagrants, et il est difficile d'obtenir des coefficients exploitables. Nous pourrions choisir un modèle à 3 variables, avec X_5 qui représenterait le PIB en euros constants mais son interprétation serait biaisé par le coté non-aléatoire de cette variable (du à son augmentation constante et linéaire). Nous jugeons qu'un bon modèle à 2 variables, qui explique plus de 50% de notre variable expliquée, serait plus adéquat, qu'un modèle qui pourrait amener à des conclusions erronées quant à l'importance de chaque variable.

IV. 2. - Interprétation du modèle définitif

IV. 2. 1. - Interprétation des coefficients

Nous optons donc pour un modèle à deux variables :

$$\hat{Y}_{(t+1)} = 0,102200 - 0,197317 * X_{3(t)} - 0,515172 * X_{4(t)}$$

Le premier coefficient (-0,197317) nous indique que si X_3 (le taux d'inflation) augmente de 1 point l'année (t), le taux de chômage de l'année (t+1) devrait diminuer de 0,2 points. Le second coefficient (-0,515172) signifie que toute augmentation d'un point de pourcentage du revenu disponible des ménages, entraîne une diminution du taux de chômage de 0,52 points.

IV. 2. 2. - Prévisions

Nous avons volontairement exclu la 36ème observation du modèle (celle de l'année 2017) pour pouvoir estimé $Y(2017)$ à l'aide de notre modèle définitif et comparer avec la vraie valeur de $Y(2017)$.

$$\hat{Y}_{(2017)} = 0,102200 - 0,197317 * 0,002 - 0,515172 * 0,0188$$

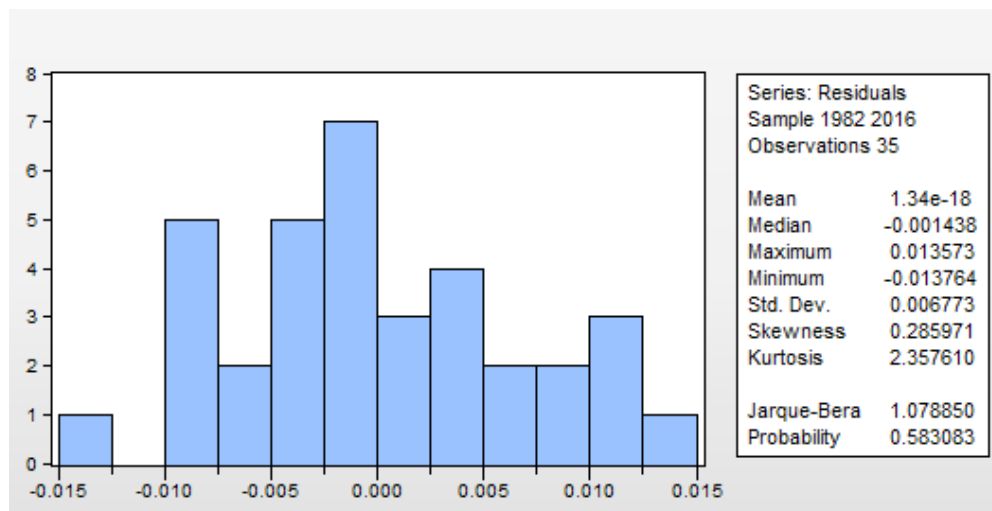
$$\hat{Y}_{(2017)} = 0,0921.$$

$$Y_{(2017)} = 0,095.$$

Notre modèle prévoit un taux de chômage de 9,27 % en 2017, tandis que la valeur observée est de 9,5 %. On peut observer un légère différence, bien sur due au fait que notre modèle ne soit constitué que de deux variables. Cependant la prévision est loin d'être mauvaise avec $Y_{(2017)}$ - $\hat{Y}_{(2017)} = 0,0029$.

IV. 2. 3. - Tests sur les résidus

- Test de normalité :



Le test de normalité nous permet d'accepter ou rejeter H_0 : les termes d'erreur sont distribués selon une loi Normale. H_0 est acceptée si la p-value de Jarque-Bera est supérieur à 0,05. C'est le cas ici puisque $0,58 > 0,05$, on peut conclure que H_0 est acceptée, les termes d'erreur sont distribués normalement.

- Test d'homoscédasticité :

Heteroskedasticity Test: White			
F-statistic	0.988165	Prob. F(5,29)	0.4419
Obs*R-squared	5.095013	Prob. Chi-Square(5)	0.4044
Scaled explained SS	2.891041	Prob. Chi-Square(5)	0.7168

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/23/19 Time: 12:18
Sample: 1982 2016
Included observations: 35

Le test de White nous permet de nous assurer que l'hypothèse d'homoscédasticité soit vérifiée. On pose H_0 : les termes d'erreur sont indépendants, cette hypothèse sera acceptée si la p-value est supérieure à 0,05. Ici la $P[X^2(5)]=0,4044>0,05$. On accepte donc H_0 , l'hypothèse d'homoscédasticité est respectée.

- Test Durbin-Watson

Ce test nous permet de vérifier l'absence d'auto corrélation des termes d'erreurs. On pose :

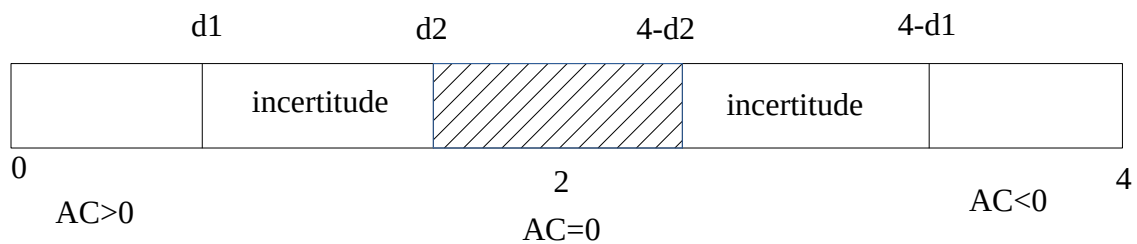
- H_0 : auto corrélation = 0
- H_1 : auto corrélation > 0
- H_1^* : auto corrélation < 0

Pour vérifier ces hypothèse nous avons besoin de construire un intervalle à l'aide de la table de Durbin-Watson :

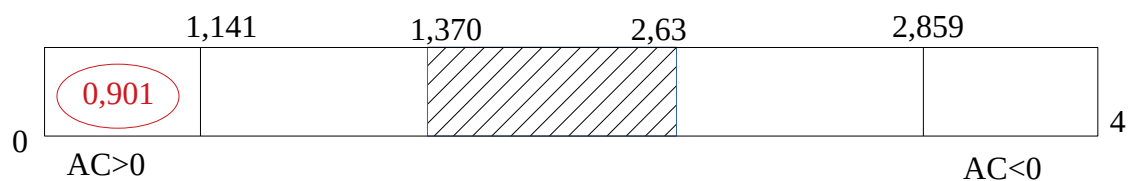
n\k	1	2	3	4	5	6	7	8	9	10
31	1.147	1.274	1.085	1.345	1.022	1.425	0.960	1.509	0.897	1.601
32	1.160	1.283	1.100	1.351	1.039	1.428	0.978	1.509	0.917	1.597
33	1.171	1.291	1.114	1.358	1.055	1.432	0.995	1.510	0.935	1.594
34	1.184	1.298	1.128	1.364	1.070	1.436	1.012	1.511	0.954	1.591
35	1.195	1.307	1.141	1.370	1.085	1.439	1.028	1.512	0.971	1.589
36	1.205	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.987	1.587
37	1.217	1.322	1.164	1.383	1.112	1.446	1.058	1.514	1.004	1.585
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.584
39	1.237	1.337	1.187	1.392	1.137	1.452	1.085	1.517	1.033	1.583
40	1.246	1.344	1.197	1.398	1.149	1.456	1.098	1.518	1.047	1.583
45	1.288	1.376	1.245	1.424	1.201	1.474	1.156	1.528	1.111	1.583
50	1.324	1.403	1.285	1.445	1.245	1.491	1.206	1.537	1.164	1.587
55	1.356	1.428	1.320	1.466	1.284	1.505	1.246	1.548	1.209	1.592
60	1.382	1.449	1.351	1.484	1.317	1.520	1.283	1.559	1.248	1.598
65	1.407	1.467	1.377	1.500	1.346	1.534	1.314	1.568	1.283	1.604
70	1.429	1.485	1.400	1.514	1.372	1.546	1.343	1.577	1.313	1.611
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.586	1.340	1.617
80	1.465	1.514	1.440	1.541	1.416	1.568	1.390	1.595	1.364	1.624
85	1.481	1.529	1.458	1.553	1.434	1.577	1.411	1.603	1.386	1.630
90	1.496	1.541	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.641
100	1.522	1.562	1.502	1.582	1.482	1.604	1.461	1.625	1.441	1.647
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725

n étant le nombre d'observations (ici 35) et k le nombre de variables explicatives du modèle (ici

2). L'intervalle se construit comme suit :



d1 étant la borne inférieure de la table et d2 la borne supérieure. Soit d1=1,141 et d2=1,370.



Dans notre modèle la statistique de Durbin-Watson vaut 0,901297. On se trouve dans la zone d'auto corrélation positive. Cela nous indique qu'il faut interpréter les résultats avec précaution, car une auto corrélation des termes d'erreur positive peut biaiser les estimateurs, notamment leur variance. Ce biais peut fausser les résultats de t-stat et l'interprétation de la significativité des variables. Cependant les tests de White et de Bera donnent des résultats satisfaisants, on peut donc décider de garder ce modèle.

Conclusion

Nous avons eu beaucoup de difficultés dans cette régression, pour plusieurs raisons. 2 variables étaient fortement corrélées, nous avons donc été contraints de partir sur un modèle à 3 variables maximum. La variable que nous avons gardé X5 est fortement corrélée à Y dans une première période qui contient les 19 premières observations de notre modèle puis cette relation disparaît dans la seconde période représentée par les 18 observations restantes. Cela a pour effet de rendre la variable X5 non-significative quand on fait fonctionner un modèle qui contient la totalité des observations. Nous avons eu l'idée de transformer la nature X5 pour surmonter cette difficulté mais la transformation (passer du taux de croissance au PIB à aux valeurs du PIB en euros constants) empêche une interprétation pertinente des résultats bien que les modèles fonctionnent bien dans ces conditions. Pour avoir un modèle adéquat nous avons décidé de supprimer X5 quitte à obtenir un modèle avec seulement 2 variables explicatives. Ce modèle a un R^2 de 0,52, ce qui signifie que nos 2 variables restantes (X3 le taux d'inflation et X4 le revenu disponible des ménages) jouent tout de même un rôle majeur dans la détermination du taux de chômage de l'année (t+1). Elles expliquent 52 % des variations de ce taux à elles deux. Les tests de normalité et d'homoscédasticité sont validés mais le test de Durbin-Watson montre une auto corrélation positive entre les termes d'erreurs, ce qui détériore la qualité de notre modèle. L'auto corrélation des termes d'erreur peut signifier que les Y sont dépendants entre eux et le modèle peut être trompeur. Cependant l'homoscédasticité étant une des hypothèses primordiales de la RLM, nous décidons de conserver le modèle.

De plus l'interprétation du modèle reste conforme à la théorie économique, il est donc fonctionnel mais à prendre avec précaution. Nous aurions pu obtenir un modèle plus efficace avec des échantillons plus larges ou en ajoutant des variables explicatives comme le taux de salaire ou le coin fiscal par exemple, qui ne figurent pas dans notre modèle définitif. Ignorer des variables explicatives importantes peut effectivement rendre notre modèle moins performant, leur effets étant consignés dans ϵ qui représente, en plus des erreurs d'estimation, les effets des variables qui ont été omises.

Pourquoi à partir des années 2000, la relation entre le taux de croissance de la production et le taux de chômage devient ambiguë ?

La fin des années 1990 et le début des années 2000 est marqué par une complexification du chômage qui ne semble plus répondre aux déterminant conjoncturels et structurels traditionnels, ce qui explique pourquoi à partir des années 2000 la relation entre le taux de croissance du PIB et la loi d'Okun ne s'applique plus dans l'économie française. La loi qui plus est ne fonctionne qu'au dessus des 2 % de croissance du PIB. En dessous la relation est beaucoup moins claire, si l'on regarde nos données, nous avons finalement peu de valeurs au dessus de 0,02 pour X_5 , car depuis les années 2000 le taux de croissance du PIB français a rarement été au dessus des 2 %. C'est la raison pour laquelle X_5 est significative sur la première période. Avant 2000 les taux de croissance sont bien plus élevés qu'après 2000 et surtout sont généralement au dessus des 2 % (4,1 % en moyenne sur la période 1982-2001). La relation que nous avons anticipée est correcte mais ne fonctionne que sur la moitié des observations qui rempli les critères nécessaires à la corrélation entre X_5 et Y . Passée ces 20 années les conditions d'application ne sont plus réunies et la loi d'Okun ne permet plus d'expliquer le taux de chômage grâce au taux de variations du PIB.

Cela peut être la conséquence du passage à l'euro en 2002 ou au phénomène de «robotisation» de l'économie qui a un impact non-négligeable sur l'emploi. Avec la révolution internet et des nouvelles technologie des années 2000, de plus en plus d'emploi sont remplacés par des robots ou de plus en plus de machines automatisables et pouvant fonctionner sans besoin d'une intervention humaine. Par exemple avec le développement du nombre d'ordinateurs dans les entreprises, le nombre d'emplois de secrétaire a considérablement chuté depuis les années 2000. Selon une étude de l'OCDE de 2016, 9% des emplois en France seraient menacé par la robotisation³⁴.

Du fait de ce phénomène, les entreprises ont pu accroître leur productivité et leur production en remplaçant des tâches effectués par des employés par des machines, ce qui d'un côté accroît le PIB et d'un autre côté contribue à augmenter le chômage. La robotisation conduit donc à une nouvelle forme de substitution du capital au travail, ce qui explique l'ambiguïté de la relation entre le taux de variation du PIB et du taux de variation du chômage en France.

34 Artz.M, Gregory.T,Ziarahn.U, The risk of Automation for jobs inOECD Countries: A Comparative Analysis,2016, consulté le 22/04/2019