BOURGINE Ilan
LAUGEL Thibault

24/05/2013
ENSAE

Relations entre le revenu et la consommation

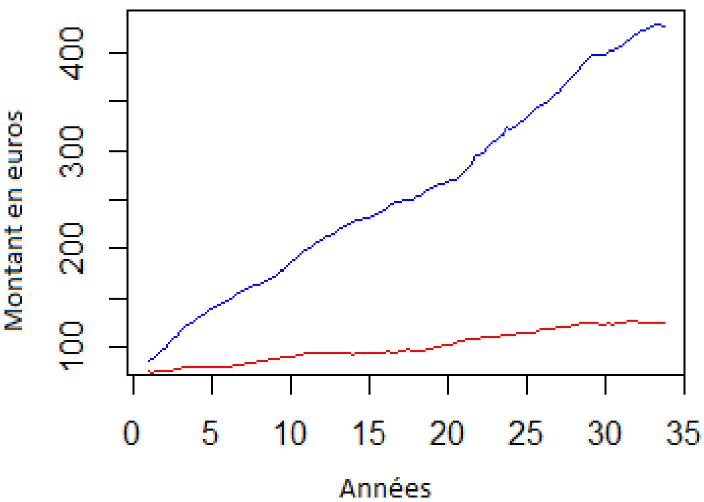
Projet de Séries Temporelles

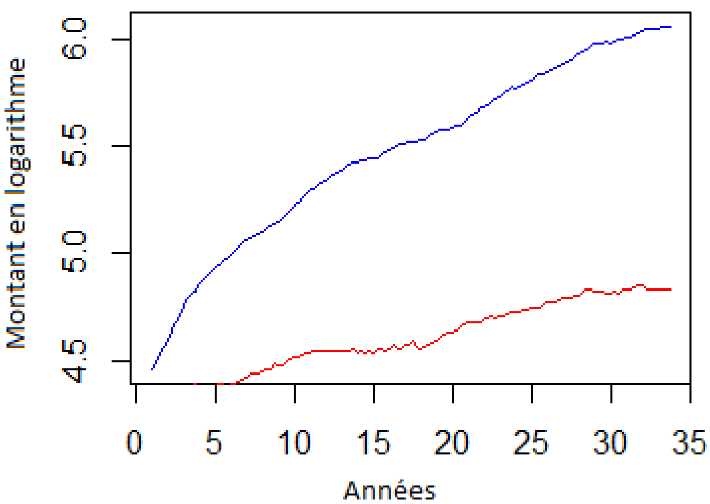
1. Les données

Les séries utilisées sont celles fournies par la Banque de Données Macro-économiques de l'INSEE. Nous avons choisi la série de la consommation mensuelle en biens des ménages en volume et celle du revenu disponible trimestrielle ajusté des ménages en prix courants. Notre série de consommation est donc en valeurs réelles tandis que notre série de revenu est en valeurs nominales, et croitra donc plus rapidement du fait de l'inflation. La consommation en biens nous a semblé un choix intéressant du fait qu'elle doit être correctement liée aux revenus disponibles puisqu'elle ne correspond pas à une nécessité comme par exemple les consommations alimentaires. Après avoir trimestrialisé la consommation, nous avons décidé d'étudier nos deux séries pour la période allant du 1er trimestre 1980 au dernier trimestre de l'année 2012, ce qui correspond à la série entière de la consommation. En gardant un grand nombre de valeurs, nous espérons obtenir des résultats les plus robustes possibles.

Nous avons représenté ci-dessous nos deux séries d'abord sans transformation logarithmique puis avec. Le revenu apparait en bleu et la consommation en rouge:

Figure 1 : Evolution des séries de la Consommation (en rouge) et du Revenu disponible (en bleu) des ménages, en valeurs et en logarithme





Plusieurs constats nous sont permis: tout d'abord, nos séries sont strictement croissantes dans le temps, donc non stationnaires, et ne présentent pas d'importante volatilité. Par ailleurs, les deux séries semblent être en partie corrélées puisque l'on observe pour les deux, par exemple, une certaine reprise de la croissance vers la 20ème année (année 2000). Finalement, la variance de nos séries semble être presque stationnaire puisque les courbes restent relativement lisses en toute période. Pour cette raison et parce que les courbes sans transformation logarithmique sont proches d'être linéaires (et donc d'être stationnaires à l'ordre 1), il ne nous semble pas judicieux d'utiliser pour le reste de l'étude la transformation logarithmique. Nous conservons donc les séries initiales.

2. Modèles ARIMA

Dans cette partie, nous aborderons la modélisation univariée d'une série à l'aide de la spécification ARIMA. En effet, afin de pouvoir correctement étudier la relation entre la consommation en biens des ménages et leur revenu disponible, nous nous sommes intéressés à l'évolution de ces séries indépendamment l'une de l'autre. La première étape de cette analyse consiste à déterminer si les séries sont stationnaires. Dans le cas contraire, nous rechercherons l'ordre auquel elles doivent **être** différenciées pour être stationnarisées. Divers tests de racines unitaires ont été utilisés pour aboutir à ce résultat : Dickey-Füller augmenté, Philipps-Perron, Elliott-Rothenberg-Stock(ERS), DFGLS et KPSS. Si tous donnent un indice sur la stationnarité des séries, les tests KPSS et ERS sont les plus puissants et seront regardés prioritairement dans toute la suite de notre étude.

Tableau 2 : Résultats des tests Dickey-Füller augmenté, Philipps-Perron et KPSS pour les deux séries

Tests	Dickey-Füller augmenté		Philipps-Perron		KPSS	
H0	Non stationnaire		Non stationnaire		Stationnaire	
Série	Consommation	Revenu disponible	Consommation	Revenu disponible	Consommation	Revenu disponible
p-value	0.749	0.8663	0.6145	0.7726	<0.01	<0.01

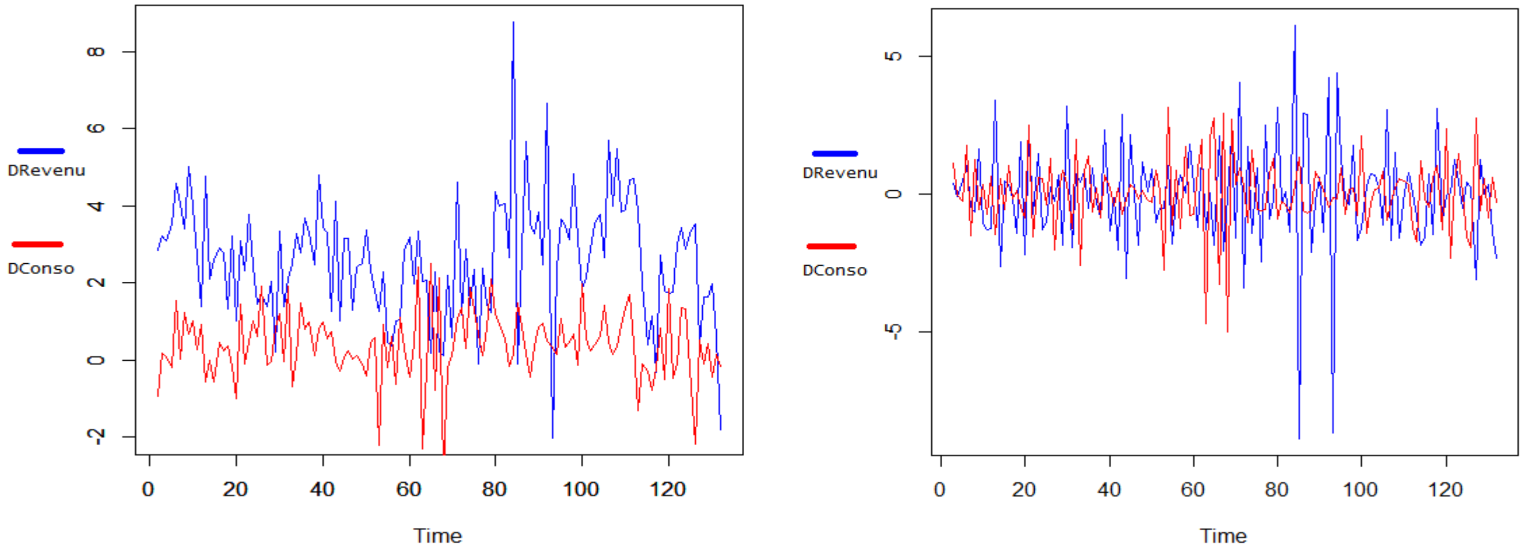
Tableau 3 : Résultats des tests DFGLS et ERS pour les deux séries

Tests	DFGLS		Elliot-Rothenberg-Stock	
H0	Non stationnaire		Non stationnaire	
Série	Consommation	Revenu disponible	Consommation	Revenu disponible
Valeur critique à 1%	-3.46	-3.46	4.05	4.05
Valeur critique à 5%	-2.93	-2.93	5.66	5.66
Valeur critique à 10%	-2.64	-2.64	6.86	6.86
Valeur de la statistique de test	-1.7248	-2.1708	17.0783	8.5649

Comme le montrent les résultats des tableaux 2 et 3 ci-dessus, l'hypothèse de stationnarité est fortement rejetée, confirmant nos premières observations de la représentation des séries. Ceci nous permet de nous faire une première idée sur la forme de la modélisation $ARIMA(p,d,q)$

recherchée : le coefficient d sera a priori non nul. Cette conclusion nous pousse également à procéder à la différenciation des deux séries.

Figure 4 : Evolution des deux séries, différenciées une fois, puis deux fois



Graphiquement, la différenciation des séries semble leur avoir fait atteindre un état stationnaire : au premier ordre pour la Consommation, au second pour le Revenu disponible. Nous allons maintenant procéder aux mêmes tests que précédemment pour conclure sur leur stationnarité. Les résultats sont donnés ci-dessous :

Tableau 5 : Résultats des tests Dickey-Füller augmenté, Philipps-Perron et KPSS pour les deux séries différenciées une fois

Tests	Dickey-Füller augmenté		Philipps-Perron		KPSS	
H0	Non stationnaire		Non stationnaire		Stationnaire	
Série	Consommation	Revenu disponible	Consommation	Revenu disponible	Consommation	Revenu disponible
p-value	<0.01	<0.01	<0.01	<0.01	>0.1	>0.1

Tableau 6 : Résultats des tests DFGLS et ERS pour les deux séries différenciées une fois

Tests	DFGLS		Elliot-Rothenberg-Stock	
H0	Non stationnaire		Non stationnaire	
Série	Consommation	Revenu disponible	Consommation	Revenu disponible
Valeur critique à 1%	-2.58	-2.58	1.91	1.91
Valeur critique à 5%	-1.94	-1.94	3.17	3.17
Valeur critique à 10%	-1.62	-1.62	4.33	4.33
Valeur de la statistique de test	-1.4921	-3.2634	0.9986	1.9464

On observe cette fois que l'hypothèse de non-stationnarité des séries est fortement rejetée par les tests Dickey-Füller augmenté et Philipps-Perron, tandis que le test KPSS met quant à lui en évidence que l'hypothèse de stationnarité des séries ne peut être rejetée. Si le test DFGLS n'est pas significatif pour la Consommation des ménages, il l'est cependant pour le Revenu disponible au seuil de 1%. Enfin, le test ERS permet de rejeter au seuil de 1% pour la Consommation et de 5% pour le

Revenu disponible l'hypothèse de non-stationnarité des séries. Choisissant, comme mentionné plus haut, de nous concentrer sur les tests KPSS et ERS, la série différenciée de Consommation des ménages semble pouvoir être considérée comme stationnaire, suggérant ainsi l'utilisation de ce premier ordre de différenciation pour sa modélisation ARIMA.

Malgré le rejet de l'hypothèse de non-stationnarité du test ERS au seuil 5% pour le Revenu disponible, nous décidons de différencier la série une deuxième fois afin d'améliorer ces résultats.

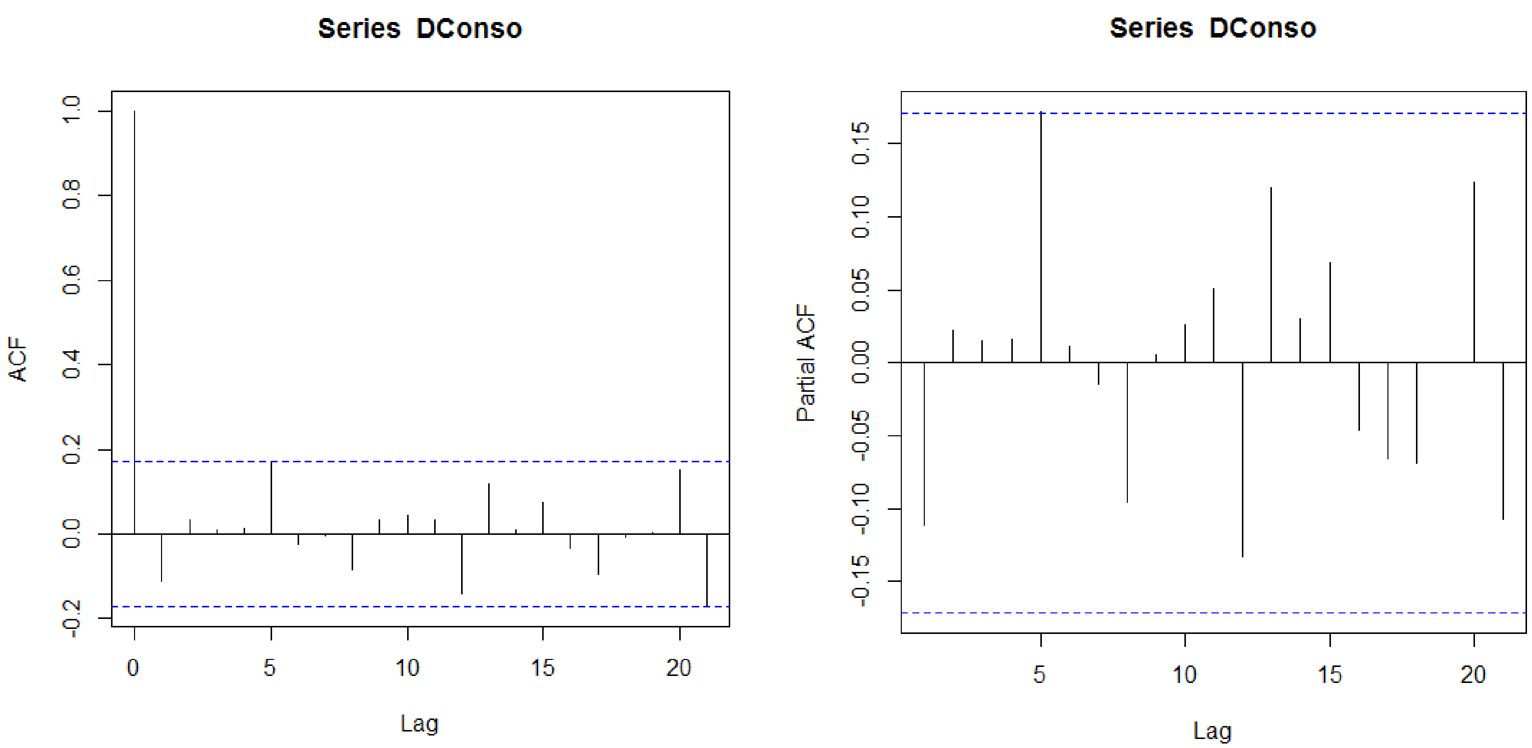
Tableau 7 : Résultats des tests DFGLS et ERS pour les deux séries différenciées une fois

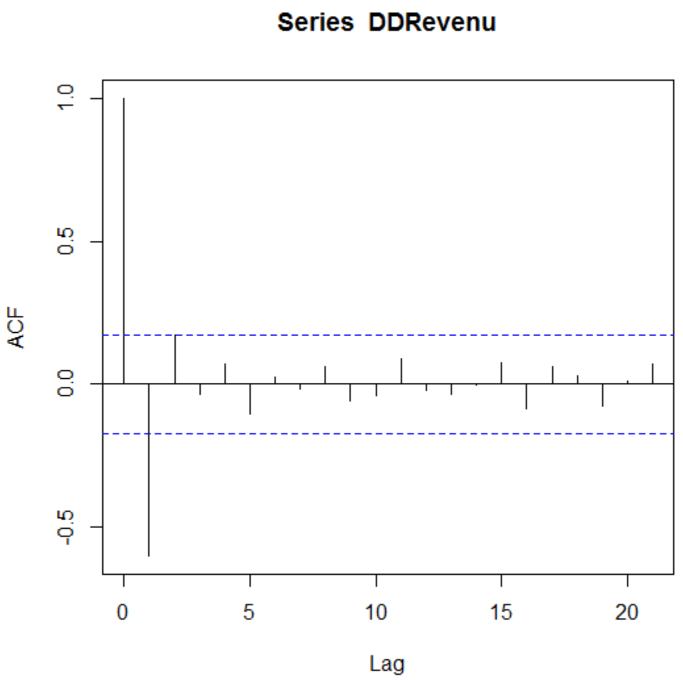
Tests	DFGLS		Elliot-Rothenberg-Stock	
H0	Non stationnaire		Non stationnaire	
Série	Consommation	Revenu disponible	Consommation	Revenu disponible
Valeur critique à 1%	-2.58	-2.58	1.91	1.91
Valeur critique à 5%	-1.94	-1.94	3.17	3.17
Valeur critique à 10%	-1.62	-1.62	4.33	4.33
Valeur de la statistique de test	-1.5971	-5.011	4.1188	0.0339

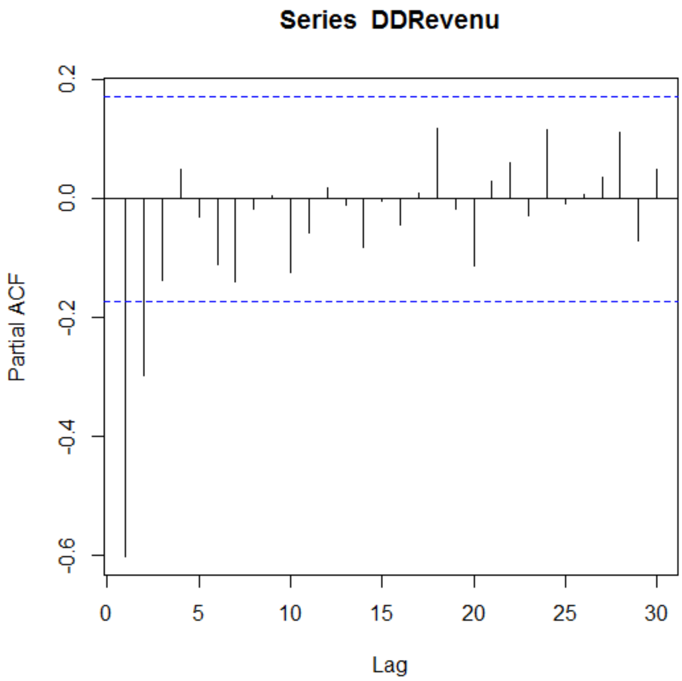
Si le test ERS n'est plus significatif pour la Consommation des ménages au seuil de 5%, il permet en revanche de rejeter l'hypothèse de non-stationnarité de la série du Revenu disponible au seuil de 1%. C'est donc au deuxième ordre que sera différenciée cette série pour sa modélisation ARIMA.

Il convient désormais de choisir l'ordre pour les parties autorégressives et moyennes mobiles. Pour ce faire, nous nous sommes penchés sur l'étude des autocorrélogrammes et autocorrélogrammes partiels des deux séries différenciées.

Figure 8 : Autocorrélogrammes et autocorrélogrammes partiels des séries de Consommation (différenciée au premier ordre) et de Revenu disponible (différenciée au second ordre) des ménages







L'observation de ces diagrammes nous permet d'avoir une idée de l'ordre des parties autorégressives et moyenne mobile. L'ordre de moyenne mobile maximal (c'est-à-dire l'ordre à partir duquel les autocorrélations ne sont plus significatives) est de 0 ou 5, tandis que l'ordre autorégressif maximal (c'est-à-dire l'ordre à partir duquel les autocorrélations partielles ne sont plus significatives) est de 5 pour la série de la Consommation en biens des ménages. Cette dernière valeur est cependant à nuancer dans la mesure où les valeurs affichées par l'autocorrélogramme partiel sont très proches de zéro pour les lags inférieurs à 5. Pour la série du Revenu disponible des ménages, ces ordres maximaux sont respectivement de 2 et 2.

Ces observations nous ont conduit à vouloir tester des modèles de type $ARIMA(p, 0, q)$ pour les séries différenciées, avec p et q compris entre 0 et 5. Afin de sélectionnerle modèle le plus robuste, nous avons utilisé les critères d'information d'Akaike (AIC), d'information d'Akaike corrigé (AICc) et le critère d'information bayésienn (BIC) : un modèle étant plus efficace qu'un autre lorsque ses critères d'informations sont inférieurs. Nous avons choisi d'utiliser ces trois critères pour comparer les résultats trouvés avec le critère AIC, le plus couramment utilisé, et ceux trouvés avec les critères BIC et AICc, qui ont l'avantage de présenter moins de sensibilité au nombre de paramètres du modèle. Ci-dessous sont présentés uniquement les résultat du calcul du AIC, les résultats pour le AICc et le BIC étant disponibles en annexe.

Figure 9 : Valeurs du critère AIC des différents modèles ARIMA pour la Consommation des ménages

AIC	q=0	q=1	q=2	q=3	q=4	q=5
p=0	364.7460	366.3447	365.0241	366.0590	365.5333	359.6533
p=1	366.2014	**350.9850**	351.0851	353.0758	355.0177	360.4629
p=2	363.7238	351.1655	353.0733	359.0081	360.9273	362.3395
p=3	364.5068	359.8708	361.5883	360.8458	363.5198	364.2883
p=4	366.3820	361.1828	362.6046	359.9218	361.7011	363.6955
p=5	368.3161	361.3704	363.2636	361.7535	363.7002	365.6896

Figure 10 : Valeurs du critère AIC des différents modèles ARIMA pour le Revenu disponible des ménages

AIC	q=0	q=1	q=2	q=3	q=4	q=5
p=0	554.1600	483.2891	483.5285	482.2460	479.5525	478.7240
p=1	497.5576	482.4755	484.2839	**477.8288**	478.9819	480.5641
p=2	485.9656	482.2521	480.4072	479.0081	480.9273	482.3395
p=3	484.5068	479.8708	481.5883	480.8458	483.5198	484.2883
p=4	486.3820	481.1828	482.6046	479.9218	481.7011	483.6955
p=5	488.3161	481.3704	483.2636	481.7535	483.7002	485.6896

La minimisation du critère AIC nous encourage à retenir le modèle (1,1,1) pour la Consommation et (1,2,3) pour le Revenu disponible. L'observation des critères bayésien BIC et d'Akaike corrigé AICc, dont les résultats sont fournis en annexe, vient renforcer ces résultats dans la mesure où le même modèle est retenu pour la Consommation. Pour le Revenu disponible en revanche, le critère d'information bayésien nous suggère la sélection du modèle (0,2,1). Les deux modèles sélectionnés par les critères ont pour ordre de la partie autorégressive 1 ou 0. Ainsi, pour chaque série, au plus seule la variable retardée d'une période interviendra dans la modélisation. L'ordre de moyenne mobile est de 1 et de 3 pour les séries de Consommation des ménages et de Revenu disponible.

Etudions maintenant le comportement des modèles sélectionnés par l'utilisation de ces critères face à certaines procédures de vérification. Nous avons tout d'abord procédé à un test de Ljung-Box, afin de tester l'autocorrélation des résidus. Les résultats, présentés en annexe, montrent qu'aucune corrélation significative n'est observable pour la série de la Consommation. La comparaison des résultats obtenus pour les modélisations (1,2,3) et (0,2,1) du Revenu disponible ne nous permet pas de sélectionner un modèle, l'ensemble des résidus n'était pas autocorrélés. Nous avons donc décidé de nous intéresser à la significativité des coefficients des régressions pour conclure.

Figure 11 : p-values des modèles ARIMA(1,2,3) et ARIMA(0,2,1) de la série du Revenu disponible

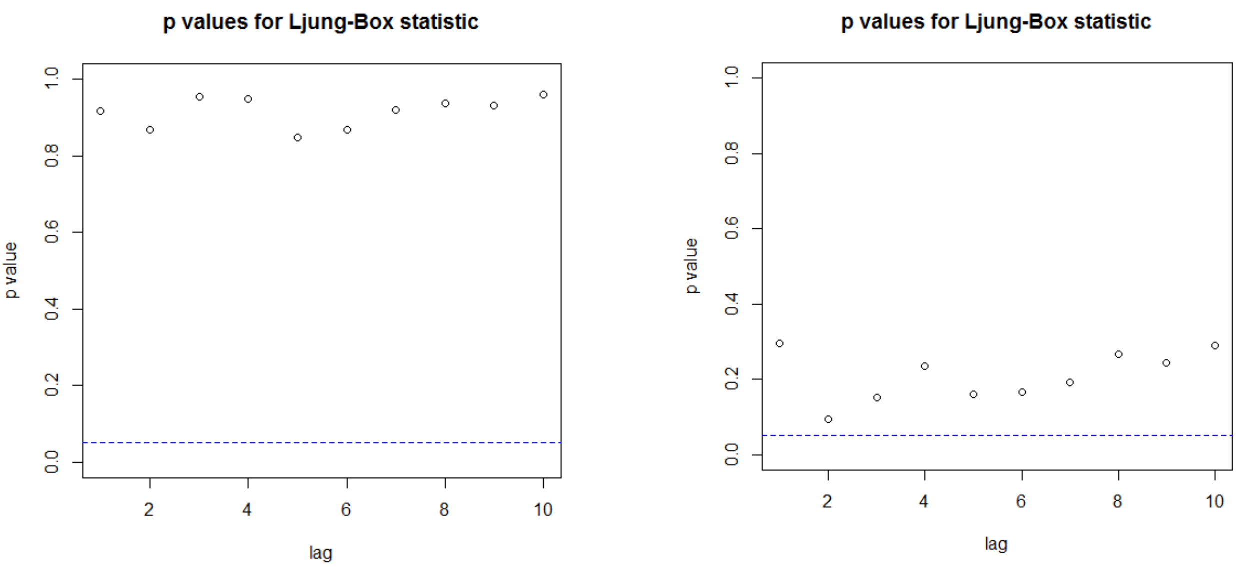


Figure 12 : Valeurs des statistiques des coefficients des régressions ARIMA de la série du Revenu

Modèle	AR1	MA1	MA2	MA3
ARIMA(1,2,3)	5.478458***	-13.138954***	5.899461***	-3.416483***
ARIMA(0,2,1)		-11.92616***		

Une fois ces valeurs de statistiques obtenues, il suffit de les comparer à celles d'une loi de Student pour évaluer leur significativité. L'ensemble de nos coefficients des deux régressions étant significatifs à 1%, nous ne sommes pas en mesure de sélectionner le modèle final via cette analyse. Nous décidons donc de nous intéresser à la normalité des coefficients.

Figure 13 : Résultats des tests de normalité des résidus des deux modélisations de la série du Revenu disponible

Modèle	p-value
ARIMA(1,2,3)	0.001227
ARIMA(0,2,1)	1.915e-11

Malheureusement, l'hypothèse de normalité des résidus est fortement rejetée dans les deux modèles, ce qui vient nuancer leur qualité et ne nous aide pas dans la sélection d'un modèle final.

3. Cointégration et modèle à correction d'erreur

Nous cherchons tout d'abord à savoir s'il existe une relation de cointégration entre nos deux séries de consommation et de revenu. Nous avons vu dans la partie précédente que ces séries différenciées une fois étaient toutes deux stationnaires (bien que le revenu différencié deux fois soit stationnaire de façon plus évidente qu'en étant différencié une unique fois). Dès lors, il ne peut exister de relation de cointégration que s'il existe une relation linéaire entre consommation et revenu non différenciés qui soit stationnaire, soit s'il existe β_1 et β_2 tels que:

$$C_t = \beta_1 + \beta_2 \times R_t + u_t$$

où la série u_t est stationnaire. Nous pouvons donc commencer par estimer les coefficients de cette régression, présentés dans le tableau ci-dessous:

Figure 14 : Coefficients de la régression simple MCO de la consommation sur le revenu

	Consommation
Revenu	0.165 (p-value < 0.0001)
Constante	58.314 (p-value < 0.0001)

Afin de savoir s'il y a une relation de cointégration, il est nécessaire d'effectuer des tests pour vérifier que le terme résiduel de cette régression est stationnaire. Nous utilisons ici un test ADS (Augmented Dickey-Fuller) dont l'hypothèse nulle est la non-stationnarité de la série. Les valeurs critiques doivent être prises à partir des tables de MacKinnon puisque la régression porte sur des résidus estimés. Dans notre modèle, les valeurs critiques à 1% et 5% sont respectivement -2.57 et -1.94. Nous trouvons pour valeur de la statistique **-2.01**, ce qui nous permet de rejeter l'hypothèse de non-stationnarité à 5% (avec une p-value de 0.045). Un test KPSS nous confirme ce résultat en ne rejetant pas l'hypothèse de stationnarité à moins de 10%.

Ces tests nous permettent de dire qu'il existe une relation de cointégration entre nos séries de revenu et de consommation. Il s'agit de la relation de long terme entre le revenu et la consommation non différenciés avec les coefficients présentés plus hauts:

$$C_t = 58.314 + 0.165 \times R_t + u_t$$

Dès lors, nous pouvons estimer un Modèle à Correction d'Erreur (MCE) intégrant les variables en variation et en niveau à l'aide de la procédure de Engle-Granger. L'emploi de ce modèle nous permettrait des prévisions plus fiables qu'en utilisant la relation de long terme puisque les résultats de l'estimation de cette relation sont faussés par la non stationnarité des séries. Ici, le Modèle à Correction d'Erreur est le suivant:

$$\Delta C_t = \alpha_1 \times \Delta R_t + \alpha_2 \times \hat{u}_{t-1} + v_t \quad avec \quad \alpha_2 < 0$$

En particulier, $\hat{u}_t$ correspond aux résidus estimés dans la précédente régression. Les coefficients calculés sont présentés dans le tableau ci-dessous:

Figure 15 : Coefficients de la régression MCO du modèle dynamique de court-terme

	Δ Consommation
Δ Revenu	0.132 (p-value < 0.0001)
Résidu estimé	-0.085 (p-value = 0.0299)

Le coefficient α_2 étant significativement négatif (au niveau 5%), notre spécification MCE est valable. Elle nous permet d'intégrer les évolutions de court terme dans un équilibre de long terme. Ce modèle mêlant long et court-terme se réécrit à l'aide des coefficients estimés précédemment:

$$\Delta C_t = 0.131 \times \Delta R_t - 0.085 \times (C_{t-1} - 0.165 \times R_{t-1} - 58.314) + v_t$$

Nous pouvons ainsi constater que les déviations par rapport à la tendance de long terme se résorbent à chaque période de 8,5%, tandis qu'à court terme, la croissance de la consommation s'explique positivement et de façon significative par la croissance du revenu. En effet, si la croissance de la consommation en T-1 est supérieure (inférieure) à celle prévue par le modèle de long terme, elle est corrigée à la période d'après d'un facteur 0.085 négatif (positif) qui permet à court terme de se rapprocher de la tendance de long terme.

D'un point de vue économique, les ménages se comportent comme s'ils voulaient lisser leur consommation en fonction des revenus espérés qu'ils envisagent de gagner. Si jamais ces revenus s'écartent de ce qu'ils espéraient gagner à une période donnée, alors ils compensent la tendance à la période suivante afin d'ajuster leur consommation.

Annexes

1. Valeurs des coefficients AICc et BIC

Consommation des ménages :

AICc	q=0	q=1	q=2	q=3	q=4	q=5
p=0	364.7770	366.4384	365.2131	366.3765	366.0133	360.3307
p=1	366.2951	**351.1740**	351.4025	353.5558	355.6952	361.3735
p=2	363.9127	351.4830	351.6943	351.8848	351.9143	352.9552
p=3	361.3371	352.3861	352.5533	352.4661	353.0748	355.0647
p=4	361.3371	352.3451	352.3021	353.3031	353.7933	353.3371
p=5	362.9727	354.3021	353.9143	353.1916	354.5114	355.7021

BIC	q=0	q=1	q=2	q=3	q=4	q=5
p=0	367.6212	372.0951	373.6497	377.5598	379.9093	376.9045
p=1	371.9518	**359.6106**	362.5859	367.4518	372.2689	380.5893
p=2	372.3494	362.6663	367.4493	371.1359	375.0407	382.7973
p=3	368.5559	367.7621	375.9457	379.5925	379.0763	388.9415
p=4	373.3990	375.9878	378.3087	379.3047	375.6701	386.0891
p=5	374.2239	379.0816	383.3749	381.0683	385.2634	389.9292

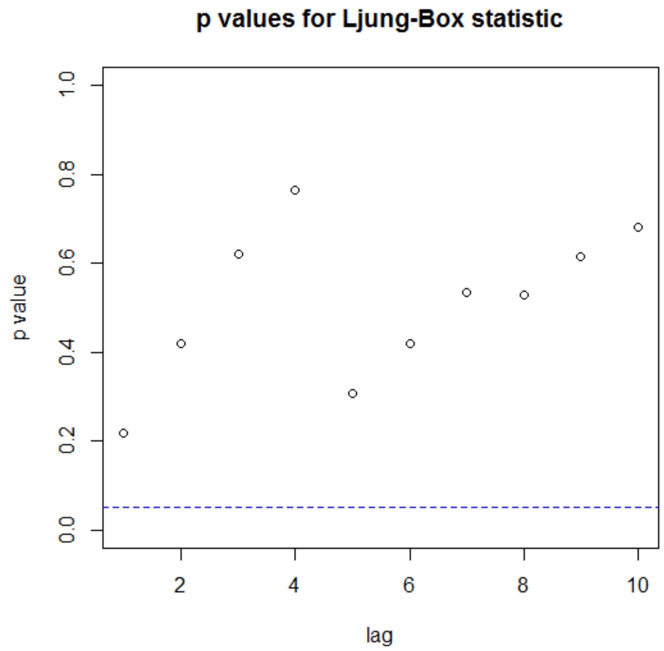
Revenu disponible :

AICc	q=0	q=1	q=2	q=3	q=4	q=5
p=0	554.1600	483.2891	483.5285	482.2460	479.5525	478.7240
p=1	497.5576	482.4755	484.2839	**477.8288**	478.9819	480.5641
p=2	485.9656	482.2521	480.4072	479.0081	480.9273	482.3395
p=3	484.5068	479.8708	481.5883	480.8458	483.5198	484.2883
p=4	486.3820	481.1828	482.6046	479.9218	481.7011	483.6955
p=5	488.3161	481.3704	483.2636	481.7535	483.7002	485.6896

AICc	q=0	q=1	q=2	q=3	q=4	q=5
p=0	555.0275	**487.0242**	490.1311	491.7162	491.8902	493.9292
p=1	501.2927	489.0781	493.7541	490.1665	494.1871	498.6369
p=2	492.5682	491.7223	492.7448	494.2133	496.9273	498.3395
p=3	494.5068	493.8708	494.5883	495.8458	496.5198	497.2883
p=4	495.3820	495.1828	494.6046	495.9218	497.7011	496.6955
p=5	496.3161	496.3704	495.2636	496.7535	498.7002	497.6896

2. Tests d'autocorrélation des résidus

Consommation



3. Code

```
library(tseries)
library(fUnitRoots)
library(stats)
library(urca)
library(forecast)

#On récupère les packages dont on aura besoin pour le projet.


base_conso_mois<-read.table("chemin vers la base consummation en biens des ménages", sep=";", header=FALSE )
base_revenu<-read.table("chemin vers la base revenu disponible", sep=";", header=FALSE )

#On a pris soin de ne garder que les mois et les trimestres que l'on souhaitait dans les bases, en enlevant les en-têtes.

base_conso_mois2 = base_conso_mois[1:394,3] + base_conso_mois[2:395,3] + base_conso_mois[3:396,3]
conso_valeur = base_conso_mois2[c(3*(0:131)+1)]
base_conso = data.frame(base_conso_mois[c(3*(0:131)+1),1], base_conso_mois[c(3*(0:131)+2),2]/3, conso_valeur)

#On crée une nouvelle base pour la consommation en sommant les 3 mois composant chaque semestre
```

```
#PARTIE 1

conso<-ts(base_conso[,3],frequency=4,names="Consommation")
revenu<-ts(base_revenu[,3],frequency=4,names="Revenu")
lconso<-ts(log(base_conso[,3]),frequency=4,names="Consommation")
lrevenu<-ts(log(base_revenu[,3]),frequency=4,names="Revenu")

#On ne garde que la série de valeurs de chaque bases, avec et sans transformation logarithmique

plot(revenu, col="blue")
lines(conso, col="red")
plot(lrevenu, col="blue")
lines(lconso, col="red")

#On crée les graphiques avec et sans transformation logarithmique faisant apparaitre revenu et consommation. On sélectionne les séries sans transformation.

#PARTIE 2

#Dickey-Füller augmenté
adfTest(revenu, type="ct")
adfTest(conso, type="ct")

#On commence par effectuer des tests de Dickey-Füller augmenté avec un modèle de régression avec constante et tendance, puisque nos séries sont croissantes et non initialement nulles.

#Philippe Perron
PP.test(revenu, lshort=TRUE)
PP.test(conso,lshort=TRUE)

#On pratique par la suite les autres tests possibles : ici Philippe-Perron.

#DFGLS
summary(ur.ers(revenu,type="DF-GLS",model="trend"))
summary(ur.ers(conso,type="DF-GLS",model="trend"))
```

```
#Test DFGLS avec tendance.


#Elliot-Rothenberg-Stock
summary(ur.ers(revenu,type="P-test",model="trend"))
summary(ur.ers(conso,type="P-test",model="trend"))

#Test ERS avec tendance.


#KPSS
kpss.test(revenu, null="T")
kpss.test(revenu, null="L")
kpss.test(conso, null="T")
kpss.test(conso, null="L")

#Test KPSS avec tendance et en niveau.

#N'ayant pas trouvé de stationnarité pour nos séries non différenciées, on recommence les même tests pour les séries différenciées une fois puis éventuellement deux fois.

____________________

drevenu<-diff(revenu)
dconso<-diff(conso)

#On différencie une fois les series.


plot(drevenu,col="blue")
lines(dconso,col="red")

#On regarde si la stationnarité semble possible.


d2revenu<-diff(drevenu)
Ddconso<-diff(dconso)

#On différencie une deuxième fois les series.


plot(d2revenu,col="blue")
lines(d2conso,col="red")

#On cherche encore à voir s'il y a stationnarité avant de reproduire les tests.
```

```
#Tests pour les séries différenciées une fois. (Cette fois, on utilise des modèles sans tendances puisqu'ils semblent être stationnaires)

adfTest(dconso,type="c")
PP.test(dconso,lshort=TRUE)
summary(ur.ers(dconso,type="DF-GLS",model="constant"))
summary(ur.ers(dconso,type="P-test",model="constant"))
kpss.test(dconso,null="L")

#Tests pour la consommation


adfTest(drevenu,type="c")
PP.test(drevenu,lshort=TRUE)
summary(ur.ers(drevenu,type="DF-GLS",model="constant"))
summary(ur.ers(drevenu,type="P-test",model="constant"))
kpss.test(drevenu,null="L")

#Tests pour le revenu.


#Tests pour les séries différenciées deux fois.

adfTest(d2conso,type="c")
PP.test(d2conso,lshort=TRUE)
summary(ur.ers(d2conso,type="DF-GLS",model="constant"))
summary(ur.ers(d2conso,type="P-test",model="constant"))
kpss.test(d2conso,null="L")

#Tests pour la consommation.




adfTest(d2revenu,type="c")
PP.test(d2revenu,lshort=TRUE)
summary(ur.ers(d2revenu,type="DF-GLS",model="constant"))
summary(ur.ers(d2revenu,type="P-test",model="constant"))
kpss.test(d2revenu,null="L")

#Tests pour le revenue.


#On a finalement décidé de conserver la série de consummation différenciée une fois et la série de revenue différenciée deux fois. On va maintenant étudier les autocorrélogrammes et autocorrélogrammes partiels.


acf(dconso)
acf(d2revenu)

#Autocorrélogrammes
```

```
pacf(dconso, lag.max=40)
pacf(d2revenu, lag.max=30)

#Aucorrélogrammes partiels


#On calcule maintenant les critères d'information.


AIC.revenu<-matrix(0,6,6)
AICC.conso<-matrix(0,6,6)
AICC.revenu<-matrix(0,6,6)
AIC.conso<-matrix(0,6,6)
for(p in 1:6){
for(q in 1:6){
ARMA.revenu<-arima(d2revenu,c(p-1,0,q-1))
ARMA.conso<-arima(dconso,c(p-1,0,q-1))
AIC.revenu[p,q]<-ARMA.revenu$aic
AIC.conso[p,q]<-ARMA.conso$aic
AICC.conso[p,q]<-AIC.conso[p,q]+2*(p+q-1)*(p+q)/(length(dconso)-p-q)
AICC.revenu[p,q]<-AIC.revenu[p,q]+2*(p+q-1)*(p+q)/(length(d2revenu)-p-q)
}}

#On obtient d'abord les critères AIC et AICc pour la consommation et le revenu.


print(AIC.revenu)
print(AIC.conso)
print(AICC.conso)
print(AICC.revenu)



BIC.fonction<-function(series,order){
model<-arima(series,order)
BIC<- -2*model$loglik + (order[1]+order[3]+1)*log(length(series) - order[2])
return(BIC)
}
BIC.revenu<-matrix(0,6,6)
BIC.conso<-matrix(0,6,6)
for(p in 1:6){
for(q in 1:6){
BIC.revenu[p,q]<-BIC.fonction(d2revenu,c(p-1,0,q-1))
BIC.conso[p,q]<-BIC.fonction(conso,c(p-1,1,q-1))
}}

#On obtient ensuite les critères BIC.

print(BIC.revenu)
print(BIC.conso)
```

```
#Afin de sélectionner le modèle, on procède à différents tests de vérification des résultats.

Aconso<-arima(conso, c(1,1,1))
Arevenu<-arima(revenu, c(1,2,3))
Arevenu2<-arima(revenu,c(0,2,1))

#test d'autocorrélation

tsdiag(Arevenu2)
tsdiag(Aconso)
residusconso<-Aconso$residuals
tsdiag(Arevenu)
residusrevenu<-Arevenu$residuals

#Test de normalité des résidus
RArevenu<-Arevenu$residuals
jarque.bera.test(RArevenu)
RArevenu2<-Arevenu2$residuals
jarque.bera.test(RArevenu2)

#PARTIE 3

reg1<-lm(conso~revenu)
summary(reg1)

#On regresse consommation par revenu

beta1<-reg1$coefficients[1]
beta2<-reg1$coefficients[2]
residu<-conso-beta1-beta2*revenu
plot(residu)

#On regarde la forme des résidus avant de tester leurs stationnarité.

adfTest(residu)
kpss.test(residu)
```

```
#On vérifie la stationnalité des résidus et l'on met en évidence une relation de cointégration.



reg2<-lm(diff(conso)~diff(revenu)+residu[1:(length(residu)-1)] + 0)
summary(reg2)

#On met en place le modèle MCE à partir de la relation de cointégration.
```