


**MINISTÈRE
DE L'EUROPE
ET DES AFFAIRES
ÉTRANGÈRES**

*Liberté
Égalité*

 **EXPERTISE
FRANCE**



**Programme de renforcement des systèmes
statistiques nationaux des pays du G5 Sahel dans
le cadre de la redevabilité de l'Alliance Sahel**

(Programme statistique G5-Sahel)

FORMATION EN EVALUATION D'IMPACT DES POLITIQUES PUBLIQUES

Formateur :

Dr. Mady OUEDRAOGO, *PhD Science de la Santé Publique* :

Orientation : Biostatistique/Démographe/ Mathématicien

Sommaire

I.	Introduction.....	3
I.1	Objectifs de la formation	3
I.2	Les objectifs spécifiques sont :	3
I.3	Résultats attendus	3
I.4	Le Dispositif pédagogique :	4
II.	Présentation des différentes méthodes d'évaluation.....	4
II.1	Inférence causale	4
II.1.1	Contrefactuels.....	4
II.1.2	Faux contrefactuels	5
II.2	Méthodes de l'évaluation d'impact.....	6
II.2.1	Assignment aléatoire : Les essais contrôlés randomisés	7
II.2.2	Les méthodes d'appariement (Matching)	8
II.2.3	Modèle de discontinuité.....	9
II.2.4	Variable instrumentale.....	10
II.3	Double différence (Diff-in-diff).....	12
II.3.1	Design de l'étude (avant-après avec groupe contrôle).....	12
II.3.2	Hypothèses fondamentales des tendances égales	13
II.3.3	Travaux pratiques	14
II.4.	Séries temporelles interrompues	30
II.4.1	Introduction	30
II.4.2	Le design des séries chronologiques interrompues	30
II.4.3	Étape 1 : un plan de série chronologique interrompu est-il approprié ?	30
II.4.4	Étape 2 : proposer le modèle d'impact	30
II.4.5	Étape 3 : analyse descriptive	31
II.4.6	Étape 4 : analyse de régression	32
II.4.6	Étape 5 : résoudre les problèmes méthodologiques	33
II.4.7	Étape 6 : vérification du modèle et analyses de sensibilité	34

I. Introduction

La statistique officielle et l'évaluation de l'impact des politiques publiques ont en commun un objectif d'éclairer la prise de décision, et elles partagent les mêmes principes communs garantissant notamment l'indépendance, la crédibilité et la transparence. Dans le cadre de la gestion des politiques publiques axées sur les résultats, la demande d'évaluation en constante évolution provient de l'administration ou des organes de contrôle externe, et elle porte sur tout ou partie du champ des politiques publiques. De ce fait, elle constitue une formidable opportunité pour valoriser les statistiques et le rôle de la statistique officielle dans les politiques officielles.

De ce fait, le rôle du statisticien dans les équipes pluridisciplinaires d'évaluation quantitative des politiques revêt d'une importance capitale pour mettre en œuvre les méthodes statistiques pour approcher rigoureusement la mesure des effets des politiques publiques. Il s'agit pour le statisticien, d'une perspective d'analyse approfondie qui suppose au préalable la collecte et la mise en forme des données statistiques documentant les différentes politiques publiques. Cette collecte fondée sur le choix d'indicateurs de résultats pertinents et fiables en fonction des objectifs de chaque évaluation, repose sur des enquêtes ou des sources administratives dont la qualité des données garantirait la mise en œuvre de procédures d'appariement pour une identification de l'impact causal des dispositifs mis en œuvre.

Le choix des méthodes et la qualité de leur réalisation dépendent de l'approche retenue par l'équipe des évaluateurs pour cerner les facteurs concomitants et mettre en avant un contre-factuel crédible. Il faut cependant noter que la qualité des évaluations d'impact dépend en grande partie de la disponibilité et du traitement statistique des données. Cette construction statistique de ce contre-factuel visant à établir l'impact causal d'une politique publique, présente des défis méthodologiques et requiert un renforcement des capacités d'analyse des statisticiens.

I.1 Objectifs de la formation

L'objectif global de cette formation est d'offrir aux participants les éléments de base pour la préparation et l'exploitation des données statistiques en vue de la production des analyses pour l'évaluation d'impact des politiques publiques.

I.2 Les objectifs spécifiques sont :

- Présenter les différentes méthodes d'évaluation d'impact d'un programme ;
- Présenter de la méthode des doubles différences ;
- Mettre en œuvre méthode des doubles différences avec des données de santé ;
- Présenter la méthode des séries temporelles interrompues ;
- Mettre en œuvre méthode des séries temporelles interrompues avec des données de santé.

I.3 Résultats attendus

- Chaque participant est capable d'utiliser la méthode des doubles différences méthodes pour évaluer d'impact d'un programme ;
- Chaque participant est capable d'utiliser la méthode des séries temporelles interrompues pour évaluer d'impact d'un programme

I.4 Le Dispositif pédagogique :

- Installation et utilisation du logiciel Stata ;
- Travaux pratiques à partir de bases de données et du « faire faire ». En effet, le formateur présente la démarche d'analyse et les participants reprennent l'exercice individuellement.

Le contenu de la formation comportait :

- Installation du logiciel STATA sur les PC individuels des participants ;
- Présentation des différentes méthodes d'évaluation d'impact des politiques de développement ;
- Mise en œuvre de la méthode des doubles différences ;
- Mise en œuvre de la méthode des séries temporelles interrompues ;
- Intérêt de l'évaluation des politiques publiques

II. Présentation des différentes méthodes d'évaluation

II.1 Inférence causale

L'inférence causale est le processus par lequel l'évaluateur tente de mesurer le lien de causalité entre une intervention et ses effets. En pratique, cela consiste à mesurer la différence entre ce qui arriverait à un individu s'il subissait une intervention et ce qui arriverait au même individu s'il ne subissait pas l'intervention. Cela est impossible, car le même individu, au même moment, ne peut pas subir l'intervention et ne pas le subir. Tout l'esprit des méthodes d'évaluation d'impact sera d'estimer cette situation contre-factuel, c'est-à-dire la situation qui allait prévaloir si l'individu n'avait pas subi l'intervention.

II.1.1 Contrefactuels

Les méthodes d'évaluation d'impact se place dans un cadre d'analyse dit « cadre de Rubin » qui systématise le recours à un contre-factuel hypothétique pour comparer la situation des bénéficiaires d'une intervention que l'on souhaite évaluer avec celle qui aurait prévalu en son absence. N'étant pas observable, le contre-factuel doit-être estimé à partir d'un groupe de contrôle qui correspond à des unités qui n'ont pas bénéficié d'une mesure, mais qui sont par ailleurs comparables aux bénéficiaires. En général, la difficulté qu'on rencontre dans l'exercice tient à la présence d'effet de sélection lié au ciblage du programme. Une hypothèse très importante du cadre de Rubin est qu'il n'existe pas d'externalités. Cela veut dire qu'une intervention sur un individu i n'a d'effet que sur la variable d'intérêt concernant cet individu. Cette hypothèse, bien que naturelle exclut certains cas : par exemple une campagne de vaccination a un impact sur la propagation de la maladie et donc sur des individus qui n'ont pas été traités directement.

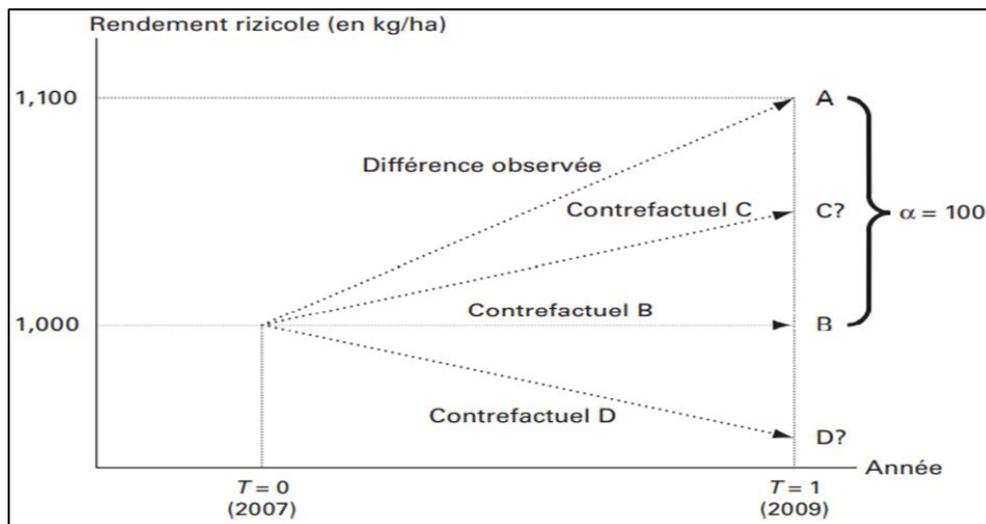
II.1.2 Faux contre-factuel

De façon intuitive, l'estimation de l'effet d'une intervention se fait en comparant une situation dans laquelle l'intervention a eu lieu à une situation dans laquelle elle n'a pas eu lieu. Ceci emmène à deux estimateurs naïfs.

II.1.2.1 Avant & après (pré & post)

Une comparaison avant-après consiste à déterminer l'impact d'un programme en examinant l'évolution des résultats pour les participants au programme au fil du temps. Autrement dit, on compare la situation des bénéficiaires avant et après l'intervention. Dans cette situation, l'estimateur a des chances d'être biaisé en raison des facteurs dits confondants ou perturbants. En effet, la comparaison repose sur l'hypothèse que si le programme n'avait pas existé, la situation pour les participants au programme aurait été exactement la même qu'avant leur participation au programme. Malheureusement, dans la grande majorité des cas, cette hypothèse n'est pas valable.

Exemple : Prenons l'exemple d'un programme de microfinance destiné aux agriculteurs pauvres en milieu rural. Ce programme propose des microcrédits aux agriculteurs pour leur permettre d'acheter des engrais afin d'accroître leur production de riz. On sait que l'année précédant le lancement du programme, la production moyenne de riz était de 1000 kg par hectare. Le programme de microfinance est lancé et l'année suivante les rendements passent à 1100 kg par hectare. Si nous cherchons à mesurer l'impact du programme en nous fondant sur une comparaison avant-après, c'est le résultat avant intervention qui constituera le contre-factuel. En appliquant la formule de base, nous concluons que le programme a permis une augmentation des rendements rizicoles de 100 kg par hectare.



Commentaire : si on imagine que les précipitations étaient normales l'année précédant la mise en œuvre du programme, mais qu'une sécheresse a lieu l'année où le programme commence. Dans cette situation, on ne peut pas considérer le résultat avant l'intervention comme le contre-factuel fiable. Puisque les agriculteurs ont bénéficié du programme lors d'une année de la sécheresse, leur rendement moyen aurait été inférieur (par exemple au niveau D) sans le programme de microfinance comme le ferait penser la comparaison avant-après. Dans cette situation, l'impact réel du programme est supérieur à 100 kg. A l'opposé, si les conditions climatiques étaient meilleures, le rendement contre-factuel aurait pu être au niveau de C et donc dans ce cas, l'impact réel du programme est inférieur à 100 kg.

II.1.2.2 Adhérents & non adhérents

La comparaison entre des unités bénéficiaires du programme et des unités n'en bénéficiant pas (« avec-sans ») constitue un autre faux contre-factuel. En effet, on compare la situation des bénéficiaires à celle de non-bénéficiaires. Ici le problème sera de trouver un bon groupe de contrôle. Deux problèmes récurrents se posent pour l'identification de ce groupe. D'abord, les programmes sont souvent ciblés vers les populations d'une certaine caractéristique ou localisé. Ensuite, la participation est au moins partiellement volontaire et dépend des caractéristiques des individus (observées ou inobservées) qui affectent également leur variable d'intérêt.

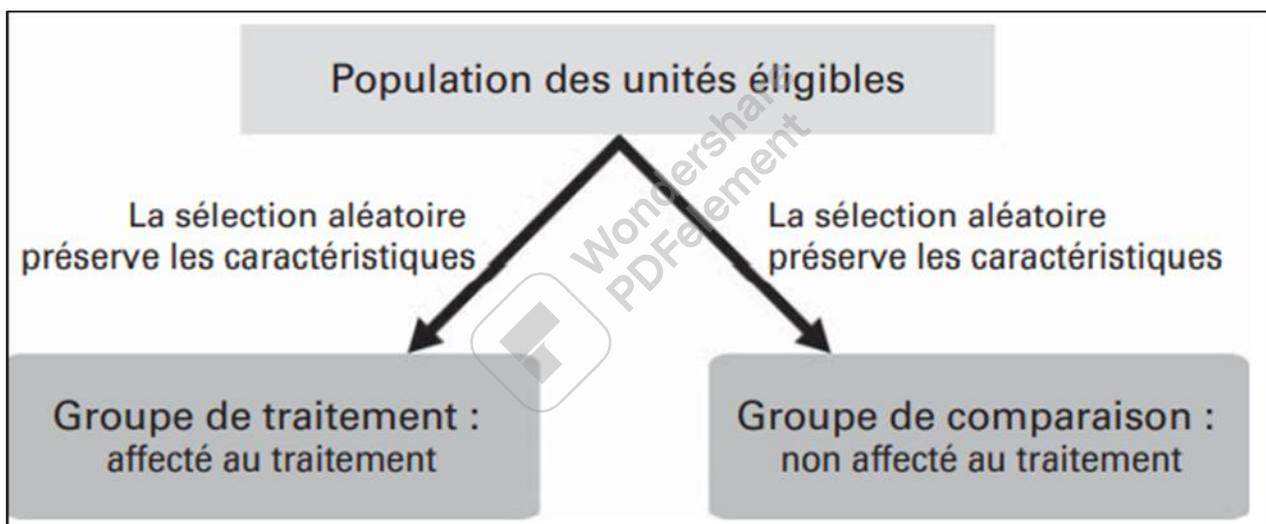
Exemple : Prenons l'exemple d'un programme de formation professionnelle destiné à des jeunes sans emploi. Imaginons que deux ans après le lancement du programme, une évaluation tente d'estimer l'impact du programme sur les revenus en comparant les revenus moyens d'un groupe de jeunes ayant participé au programme aux revenus de ceux qui n'y ont pas participé. Supposons que les jeunes ayant participé au programme aient un revenu deux fois supérieur à celui des jeunes n'ayant pas participé. Ce résultat est un résultat biaisé. En effet, autant les adhérents ont une motivation, les non adhérents en ont une. Ainsi, le fait que les individus ne préfèrent pas participer au programme de formation pour une motivation quelconque introduit un biais dans l'estimation de l'impact du programme. Ce biais est appelé biais de sélection.

II.2 Méthodes de l'évaluation d'impact

Les méthodes empiriques d'évaluation d'impact ont été développées pour estimer l'effet propre à un programme afin d'avoir des conclusions sur son efficacité. On distingue les méthodes expérimentales des méthodes naturelles ou quasi-expérimentales.

II.2.1 Assignation aléatoire : Les essais contrôlés randomisés

Les expérimentations contrôlées randomisées consistent à tester plusieurs programmes publics sur des individus qui sont désignés comme « bénéficiaires ». Ils sont comparés à des individus « témoins » qui sont intégrés dans le processus expérimental. Les individus sont assignés de façon aléatoire, aussi bien dans le groupe des bénéficiaires que dans le groupe de contrôle. Autrement, les expérimentations contrôlées consistent à tirer au hasard les personnes qui vont bénéficier du programme dont on souhaite évaluer l'efficacité et d'autres qui n'en bénéficieront pas. Dans la théorie statistique de la conception des expériences, la randomisation consiste à répartir au hasard les unités expérimentales entre les groupes de traitement. L'utilisation de l'assignation aléatoire pour définir le groupe de traitement et le groupe de comparaison garantit en théorie que les groupes sont équivalents autant sur les observables que sur les inobservables. La seule différence est l'un des groupes bénéficiera du traitement. D'où une meilleure estimation du contre-factuel.



Commentaire : une fois la population éligible est arrêtée, on procède au tirage des groupes de traitement et de comparaison. La sélection aléatoire permettra de préserver les caractéristiques la population aussi bien dans le groupe de traitement que dans le groupe de comparaison. Imaginons que la taille de la population des unités éligibles est de 2000 personnes. Imaginons que les proportions des « jeunes » et de « chômeurs » sont respectivement de 60% et de 20% s'agissant des caractéristiques observables. Pour ce qui est inobservable, imaginons par exemple que 40% ont un « QI » très élevé. Supposons maintenant que l'on doit tirer 200 personnes pour le groupe de traitement et 200 unités pour le groupe de comparaison. La force du hasard fera qu'on aura à peu près les mêmes proportions du point de

vue des caractéristiques aussi bien dans le groupe de traitement que dans le groupe de comparaison.

Alors que le modèle causal de Rubin fournit un cadre, définir les paramètres causaux (c'est-à-dire les effets d'un traitement randomisé sur un résultat), l'analyse des expériences peut prendre un certain nombre de formes. Le plus souvent, les expériences randomisées sont analysées à l'aide d'une ANOVA, d'un test t de student, d'une analyse de régression d'un test statistique.

L'assignation aléatoire constitue une méthode d'évaluation d'impact adéquate dans deux cas fréquents :

- Si le nombre d'unités éligibles est supérieur au nombre de places disponibles dans le programme. Si la demande dépasse l'offre, un tirage au sort peut être effectué pour définir le groupe qui bénéficiera du programme parmi la population éligible. Dans ce cas, toutes les unités de la population ont la même chance d'être sélectionnées.
- Lorsqu'un programme doit être progressivement étendu pour couvrir l'ensemble de la population éligible. Quand un programme est graduellement mis en œuvre, la sélection aléatoire de l'ordre dans lequel les participants bénéficieront du programme donne à chaque unité éligible une chance égale de recevoir le traitement à la première phase ou à une phase ultérieure du programme.

Les expériences contrôlées randomisées nécessitent une intervention en amont de la mise en place de la politique qu'il s'agit d'évaluer. Pourtant, le statisticien est plutôt très souvent appelé à estimer l'impact de politiques déjà mis en place.

Plusieurs autres méthodes ont été développées pour ces cas. Nous avons les méthodes appariement et les méthodes dites naturelles ou quasi-expérimentales.

II.2.2 Les méthodes d'appariement (Matching)

Ce que le statisticien peut faire pour réduire le biais de sélection, est au moins de contrôler les différences observables entre les bénéficiaires et les non bénéficiaires. Les méthodes d'appariement consistent à comparer chaque personne traitée à un « jumeau » non traité possédant les mêmes caractéristiques. Étant donné que pour un bénéficiaire i , il est très rare de lui trouver un « jumeau » dans le groupe contrôle, l'estimation de l'effet moyen de la mesure sur les individus traités est faite en appariant chaque personne de l'échantillon des bénéficiaires dans l'échantillon des non bénéficiaires en utilisant les techniques des plus proches voisins, les scores de propension ou encore la technique de l'appariement par

noyau. On fait généralement recours à ces méthodes lorsqu'il n'est pas possible de disposer de sources d'identification exogènes. Ces méthodes supposent que conditionnellement aux caractéristiques observables, le fait de choisir de suivre ou non le traitement est indépendant des caractéristiques inobservables. C'est l'hypothèse **d'indépendance conditionnelle**.

Prenons deux individus avec des caractéristiques observables identiques, l'un est traité et l'autre pas. L'hypothèse d'indépendance conditionnelle signifie que l'assignation dans un programme d'individus ayant les mêmes caractéristiques X est indépendante de leurs résultats potentiels avec et sans le programme. Elle signifie simplement qu'en dehors des caractéristiques observables, il n'y a pas d'autres caractéristiques qui influencent à la fois les résultats potentiels et l'adoption du programme dont on souhaite évaluer l'impact.

Les techniques des plus proches voisins : la méthode la plus simple est de comparer chaque individu avec son jumeau. Cela se fait en utilisant une métrique. Cette dernière peut être la distance euclidienne où la distance entre deux individus est la somme de la distance de tous les covariables.

Les scores de propension : si l'hypothèse d'indépendance conditionnelle est vérifiée, alors les résultats potentiels sont aussi indépendants de l'adhésion au traitement conditionnellement au score de propension.

Appariement par noyau : le contre-factuel de l'unité i est calculé en utilisant une moyenne pondérée de toutes les unités non traitées.

Les méthodes d'appariement impliquent d'avoir suffisamment de variables pour parvenir à cette condition. Ces méthodes sont de moins en moins utilisées, car qui l'hypothèse qui les sous-tend semble dans beaucoup de cas assez peu plausible.

Contrairement aux méthodes d'appariement, les méthodes naturelles utilisent les variations exogènes de l'environnement économique qui créent presque une situation expérimentale. Nous les présentons dans la suite.

II.2.3 Modèle de discontinuité

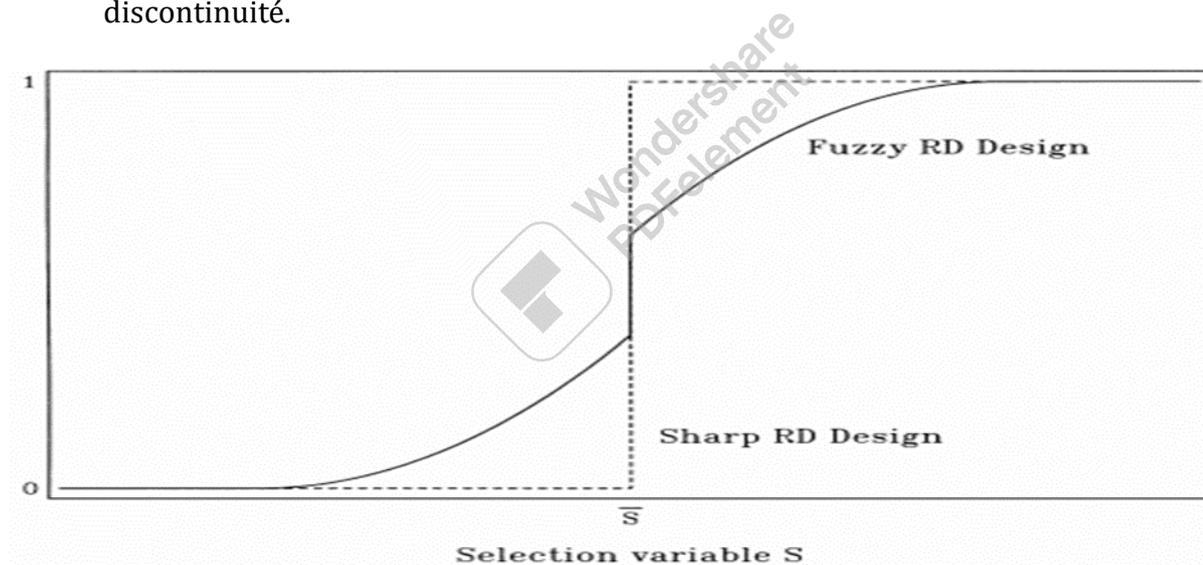
Le modèle de discontinuité repose sur l'existence d'une variable de sélection ayant un impact discontinu sur la probabilité de bénéficier ou non d'un traitement. Plus précisément, la méthode consiste à exploiter le fait que de nombreux programmes utilisent des effets à seuil (exemple : allocation sous condition de ressources). L'hypothèse est que les personnes autour de ces seuils ont des caractéristiques inobservables identiques et donc que pour ces personnes

dont les caractéristiques sont très proches, le fait d'être sélectionné peut-être considéré au premier ordre comme indépendant des caractéristiques des individus. Il est donc possible d'identifier l'effet causal du traitement en comparant les unités situées autour des seuils d'éligibilité.

Exemples d'intervention présentent des règles d'affectation à seuils : allocations sociales sous condition de ressources, contrainte d'âge pour certains dispositifs...

En pratique, deux cas doivent être distingués :

- Cas du **Sharp design** : les protocoles sont avec des discontinuités franches. Le traitement dépend de manière déterministe de la variable de sélection. Autrement dit, les unités situées de part et d'autre du seuil sont soit traitées ou pas du tout traitées.
- Cas du **Fuzzy design** : les protocoles sont avec des discontinuités floues. La variable de sélection n'affecte plus parfaitement la probabilité d'être traité. Cette fois, c'est la probabilité d'être traité conditionnellement à la variable de sélection qui présente une discontinuité.



Au niveau du sharp Design, la discontinuité est observée du point de vue de la variable de sélection alors que pour le fuzzy Design la discontinuité est observée au niveau de la probabilité d'être traité.

II.2.4. Variable instrumentale

La mise en œuvre d'une méthode instrumentale pour évaluer l'impact d'une intervention requiert de disposer d'un instrument, c'est-à-dire d'une variable qui explique le fait que certains individus aient une plus grande chance d'être affectés par cette intervention que d'autres, mais qui n'a pas d'impact direct sur les résultats de cette intervention. Donc l'hypothèse est qu'il

existe des variables qui expliquent le fait d'être traité, mais qui ne sont pas corrélées aux composantes inobservées du revenu, et qui permettent donc de s'affranchir du problème de sélection.

Pour être considérée comme une variable instrumentale valable, une variable doit posséder deux caractéristiques :

- Elle doit être corrélée avec la participation au programme,
- Elle ne doit pas être corrélée avec les résultats Y (sauf à travers la participation au programme), ni avec les variables non observées.

Si une telle variable existe, il est possible d'obtenir une estimation convergente du paramètre d'intérêt. Cependant, elle mesure les effets de la mesure en moyenne pour une population particulière, celle qui réagit à l'instrument. Cela dit, les résultats obtenus doivent être interprétés en tenant compte du fait qu'il s'agit d'une estimation locale.



II.3 Double différence (Diff-in-diff)

II.3.1 Design de l'étude (avant-après avec groupe contrôle)

On dispose de deux groupes observés avant et après l'introduction de la mesure à évaluer. L'intervention ne concerne que l'un des deux groupes. La méthode consiste à comparer l'évolution de la variable d'impact des bénéficiaires avant et après le traitement à celle des personnes du groupe de contrôle sur la même période. Elle combine les estimateurs biaisés vus plus haut pour estimer un meilleur contre-factuel. La différence dans les résultats avant-après pour le groupe participant contrôle pour les facteurs invariables dans le temps qui affectent ce groupe, pour la simple raison que nous comparons le groupe à lui-même. La différence avant-après ne tient toutefois pas compte des facteurs externes variables dans le temps. Pour prendre en compte ces facteurs externes variables dans le temps, il s'agira de mesurer la différence de résultats avant-après pour un groupe qui n'a pas participé au programme, mais qui a été exposé aux mêmes conditions externes. La méthode de double différence est une technique qui soustrait la différence de l'indicateur d'impact entre les deux groupes à la situation de référence à la différence de l'indicateur d'impact entre les deux groupes aux moments choisis pour l'évaluation d'impact. Ainsi, l'utilisation de la méthode de double différence nécessite de disposer de données de référence.

Formalisation

Supposons que les individus sont observés à deux dates, 0 et 1.

En $t = 0$ le programme n'existe pas.

En $t = 1$ le programme est en place.

On note Y_{it} la variable d'intérêt observée à la date t .

L'estimateur de la double différence s'écrit :

$$DD = E(Y_{i1} - Y_{i0} | T_i = 1) - E(Y_{i1} - Y_{i0} | T_i = 0)$$

Si l'estimateur de la double différence est calculé par la régression linéaire, alors le modèle à estimer est de la forme :

$$Y_{it} = \alpha + \beta T_i * t + \rho T_i + \gamma t + \theta X_{it} + \varepsilon_{it}$$

β est le terme d'interaction entre le temps t et la variable de traitement T_i . Il correspond à l'estimateur de la double différence.

Les variables t et T_i sont incluses séparément pour tenir compte d'un effet potentiel du temps qui passe (trend) et d'un effet provenant du fait d'être traité ou non.

X_{it} vecteur de covariables.

II.3.2. Hypothèses fondamentales des tendances égales

La méthode de la double différence permet de tenir compte des différences invariables entre le groupe de traitement et le groupe de comparaison dans le temps. Cependant, elle ne permet pas d'éliminer les différences variables entre ces deux groupes au cours du temps. Par conséquent, pour que la méthode puisse fournir une estimation valable du contre-factuel, il est nécessaire de faire l'hypothèse qu'il n'y a pas de différences variables dans le temps entre le groupe participant et le groupe contrôle. C'est l'hypothèse de trajectoires parallèles. Les groupes des traités et contrôlés doivent avoir, du point de vue des variables d'impact, une évolution tendancielle similaire.

Comment vérifier la vraisemblance de cette hypothèse ?

- Comparer les évolutions passées des variables d'impact : s'assurer que sur les périodes antérieures, les groupes des traités et de contrôle ont connu une évolution similaire. Cela nécessite d'avoir de données sur au moins deux points dans le temps autour d'une période où le traitement n'a pas eu lieu. Un graphique des évolutions de la variable d'impact observées pour les différents groupes peut permettre de se faire une idée sur la crédibilité de l'hypothèse identifiante.
- Tests « placebo » : si on dispose des données antérieures des groupes des traités et de contrôle, on procède à la mise en œuvre de la méthode de double différence autour d'une date où il n'y a eu aucune intervention. Si on note une différence significative de cette mesure placebo, on peut suspecter une différence de trajectoire entre ces deux groupes, et donc le groupe contrôle n'est pas un contre-factuel valable.
- **Triples différences** : lorsqu'on dispose de plusieurs groupes de contrôle possibles, il est judicieux de vérifier que l'effet mesuré est identique selon qu'on utilise l'un ou l'autre de ces groupes.

En résumé, évaluer l'impact d'une mesure par la méthode des doubles différences requiert :

- Choisir un groupe de contrôle : il ne sera crédible que si on peut supposer que les évolutions des revenus sont les mêmes que celles qu'auraient connues les bénéficiaires en l'absence de la mesure.

- Disposer de données cohérentes avant et après la mise en place de la mesure est indispensable. Il est préférable que les données soient relativement proches dans le temps, pour éviter de capter d'autres évolutions qui ne seraient pas liées à la mesure.

II.3.3 Travaux pratiques

Titre : Évaluation de l'effet de la chimio prévention du paludisme saisonnier sur le fardeau du paludisme chez les enfants de moins de 5 ans au Burkina Faso.

II.3.3.1 Introduction

En 2014, le gouvernement du Burkina Faso a lancé le programme de chimioprévention saisonnière du paludisme (CPS). Le bénéfice attendu était une réduction de 75 % de tous les épisodes de paludisme et une baisse de 75 % des épisodes de paludisme grave. Cette étude a évalué l'efficacité de la CPS sur la morbidité palustre dans le pays après 2 ans de mise en œuvre.

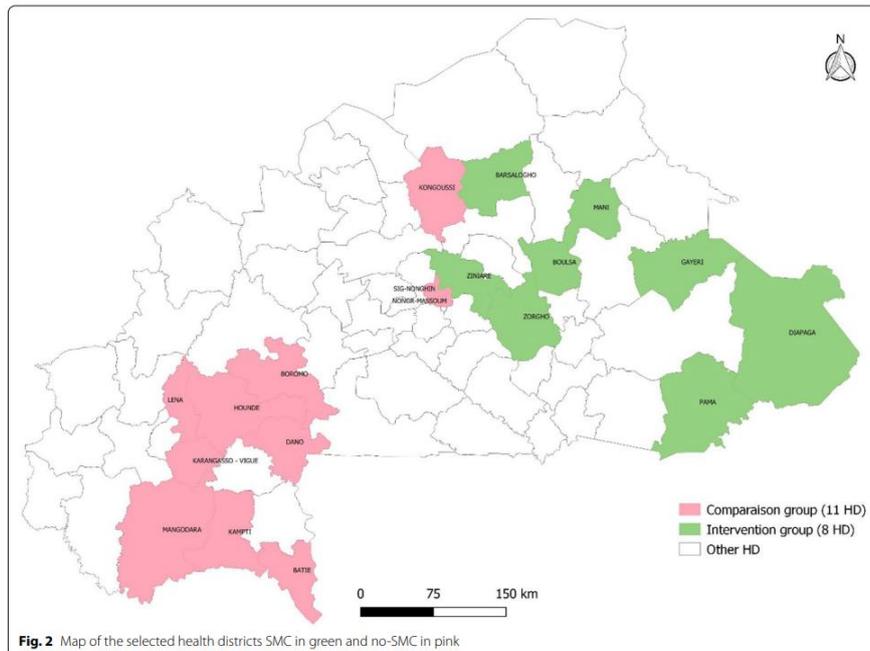
II.3.3.2 Méthodes

Conception de l'étude

Conception quasi-expérimentale comparant les changements de résultats au cours de la période de transmission élevée (août-novembre) entre les districts sanitaires SMC et non-SMC avant (2013-2014) et après l'intervention (deux cycles en 2015 et 2016). Les indicateurs de santé (nombre de cas de paludisme simple et de cas de paludisme grave de 16 districts de santé (8 en intervention et 8 dans le groupe de comparaison) ont été extraits de la plate-forme basée sur le système d'information sanitaire du district (DHIS2) comprenant les données les établissements de santé. L'effet sur l'incidence a été évalué en ajustant le modèle (DIFF and DIFF) de régression à effets mixtes en utilisant une distribution binomiale négative à une échelle logarithmique.

- **Sélection des districts de santé SMC et non-SMC**

En 2013, l'annuaire des statistiques sanitaires recensait 70 districts sanitaires (données formations sanitaires primaires et hôpitaux de district). Les critères d'inclusion pour le groupe **d'intervention étaient la CPS mise en place en 2015 et 2016** et pour le groupe de comparaison, **pas de CPS avant 2017**. Le critère d'exclusion était la présence d'un hôpital régional ou universitaire dans le district pour les deux groupes. Les districts sanitaires avec de tels hôpitaux ont été exclus parce que leurs données rapportées ne discriminent pas les patients des districts SMC et districts non SMC. Le district sanitaire de Diébougou dans lequel l'IRS a été menée pendant 3 années consécutives à partir de 2010 a été exclu.



- **Sources des données**

Les données ont été extraites du référentiel national de données sanitaires en ligne qui est une plate-forme basée sur le système d'information sanitaire de district (DHIS2) appelé ENDOS-BF (Entrepôt National de Données sanitaires du Burkina Faso), opérationnelle depuis 2013. Les rapports mensuels de données sur la santé primaire et des activités des formations sanitaires et des hôpitaux de district envoyées au district de santé, sont agrégées dans ENDOS-BF.

- **Sélection des variables d'étude**

Parmi les bénéfices attendus de la CPS, il y avait une baisse de 75 % de l'incidence de tous les épisodes de paludisme et une baisse de 75 % de l'incidence des épisodes de paludisme grave. L'étude a examiné l'effet de la CPS sur la morbidité palustre en se basant sur deux variables de résultats collectées auprès des établissements de santé primaires et des hôpitaux de district pour se conformer à la politique nationale de définition des cas de paludisme : l'incidence du paludisme non compliqué (UM), c'est-à-dire tous les cas présumés basés sur les données des symptômes cliniques ou cas parasitologiques confirmés d'un paludisme tests de diagnostic rapide (TDR) ou microscopie et incidence des cas de paludisme grave chez les enfants, c'est-à-dire cas confirmés de paludisme avec des signes de maladie grave et/ou des signes de dysfonctionnement d'organes vitaux qui ont nécessité une hospitalisation. L'incidence du paludisme simple chez les enfants de moins de cinq ans est définie comme le nombre de cas de paludisme simple rapporté à la population totale d'enfants de moins de cinq ans vivant dans le district sanitaire. L'incidence du paludisme grave chez les enfants de moins de cinq ans est définie comme le nombre de cas de paludisme grave signalés, par rapport à la population totale d'enfants de moins de cinq ans vivant dans le district sanitaire.

- **Assurance qualité des données**

Des audits de la qualité des données sont menés périodiquement dans les formations sanitaires primaires publiques par des équipes de supervision au niveau du district pour vérifier la

cohérence entre les données sources (registres) et les données rapportées dans les rapports mensuels d'activités. Par ailleurs, parallèlement aux rapports mensuels d'activité obligatoires, chaque formation sanitaire produit un télégramme hebdomadaire sur les principaux indicateurs de santé parmi lesquels figurent les cas de paludisme simple et grave chez les enfants de moins de cinq ans. Les chiffres communiqués à la Direction des statistiques sanitaires du ministère de la Santé doivent correspondre à ceux des télégrammes hebdomadaires agrégés communiqués à la direction de la surveillance des maladies du même ministère. ENDOS-BF a un module d'évaluation de la qualité intégrée, qui permet aux gestionnaires de données centraux de vérifier plusieurs dimensions de la qualité au cours de la saisie des données et de générer des requêtes destinées aux gestionnaires de données des districts. Une évaluation d'ENDOS-BF en 2017 a rapporté l'exactitude des données, l'actualité des données, l'exhaustivité des données et la fiabilité des données de 83 %, 80 %, 78 % et 87 %, respectivement. Par ailleurs, l'indice d'écart des données entre la source et les données rapportées s'est amélioré de 61 % en 2013 à 80 % en 2018.

- *L'analyse des données*

Caractéristiques contextuelles des districts sanitaires dans les groupes d'intervention et de comparaison

Les districts sanitaires de l'étude étaient comparables à bien des égards au départ (2013). Les différences de superficie sont imputables à un quartier significativement plus large (Diapaga) dans le groupe d'intervention. Il n'y avait pas de différence statistique en termes de précipitations annuelles moyennes et de température annuelle moyenne (chevauchement de l'intervalle de confiance à 95%). Le ratio population par agent de santé, dans les districts d'intervention est inférieur par rapport aux districts de comparaison. La proportion de la population vivant à moins de 5 km d'une formation sanitaire est également sensiblement plus faible dans les districts d'intervention. Globalement, la possession de moustiquaires était inférieure dans le groupe SMC par rapport au groupe non SMC en 2014, et la proportion d'utilisation de moustiquaires est similaire dans les deux groupes avant la mise en œuvre du SMC (tableau 1).

Table 1 Summary of study health districts characteristics in 2013

Background characteristics	SMC health districts (intervention)	Non-SMC health districts (comparison)
Total surface area (km ²)	46,785	33,833
Average annual rainfall (mm) (95% CI)	794 (640, 948)	752 (629, 876)
Average monthly temperature (°C) (95% CI)	28.1 (27.7, 28.6)	28.0 (27.6, 28.4)
Total population	2,042,236	2,133,367
Total number of primary health facilities	212	199
Total health staff	1150	2173
Population/health facility ratio	9633	10,720
Population/health care worker ratio	1776	981
Population of children under five		
Total	422,272	382,191
Proportion	20.7%	17.9%
Population at < 5 km away from the health facility		
Total	808,066	1,363,340
Proportion	39.6%	63.9%
Insecticide treated nets in 2014 ^a		
Ownership	75.4%	82.2%
Use	88.0%	87.2%

^a We obtained this information from the Malaria Indicator Survey (MIS) conducted in 2014. The MIS was not designed to provide estimates at health districts level. To address this issue, we fitted a Bayesian binomial model to the aggregated data on ITN ownership and use to estimate them at the district level

- ***Caractéristiques des districts sanitaires en intervention et groupes de comparaison***

Analyses descriptives et de régression

Pour chaque résultat (incidence du paludisme simple et incidence du paludisme grave) et chaque groupe, la différence moyenne entre le post-(combinant les données du premier tour du 1er août au 30 novembre 2015 et du second tour du 1er août au 30 novembre 2016) et les périodes pré-intervention (du 1er août au 30 novembre 2013 et du 1er août au 30 novembre 2014) ont été calculées comme précédemment. Le changement moyen a ensuite été soustrait des résultats au fil du temps dans les districts de santé ne mettant pas en œuvre la CPS de ceux des districts de santé mettant en œuvre la CPS pour isoler l'effet de la CPS. L'hypothèse de tendance parallèle guide l'utilisation de l'analyse de la différence de différence (DID) et suppose qu'il n'existe aucune différence variant dans le temps entre le CMS et les groupes de comparaison. À cette fin, un test a été effectué pour vérifier si les variations des résultats étaient équivalentes entre les deux groupes au cours de la période pré-intervention. En raison de la surdispersion des données, un modèle de régression binomiale négative à effets mixtes a été ajusté. Compte tenu de la variation géographique de l'incidence du paludisme, les districts de santé ont été inclus en tant qu'intersections aléatoires.

En 2016, le Burkina Faso a introduit la prise en charge gratuite des enfants de moins de 5 ans et des femmes enceintes dans les centres de santé. Cela a été associé à une augmentation marquée de la fréquentation des consultations externes et donc à une augmentation marquée des cas signalés de paludisme. Pour cela, nous avons considéré la variable comme une variable de contrôle dans l'analyse des doubles différences pour estimer avec précision l'impact de la CPS.

La validité de la régression a été incluse en comparant le critère d'information d'Akaike (AIC) de la régression de Poisson et la régression binomiale négative. Avec l'AIC le plus bas, un modèle

de régression binomiale négative à effets mixtes a été ajusté pour estimer l'effet de la CPS sur l'incidence des cas de paludisme non compliqué et des cas de paludisme grave. Chaque modèle de régression incluait une variable temporelle (période pré-SMC vs période post-SMC), une variable SMC (groupe SMC vs groupe non-SMC), la variable de soins de santé gratuits (Free_Healthcare) et la taille de la population (enfants de moins de cinq ans dans chaque district) comme un décalage.

Le modèle empirique estimé peut être caractérisé comme :

$$Y_{it} = \beta_{0i} + \beta_1 Time + \beta_2 SMC + \beta_3 Free_Healthcare + \beta_4 SMC * Time + \varepsilon_{it}$$

où Y_{it} est le nombre de cas d'intérêt fournis par le district i au cours de la période t (en mois), le temps est une variable capturant la période pré (2013-2014) et post-CPS (2015 et 2016), la CPS est un indicateur pour un district de santé faisant partie d'un groupe SMC ou d'un groupe non-SMC, Free_healthcare variable pour la gratuité des soins de santé, $SMC * Time$ l'interaction qui capture l'estimateur DID à l'échelle logarithmique, l'effet de la mise en œuvre de deux cycles de SMC en 2015 et 2016 et ε_{it} le terme d'erreur.

À partir des modèles de régression binomiale négative à effets mixtes, le rapport d'incidence, le rapport des taux avec un intervalle de confiance à 95 % et les valeurs de p ont été estimés. La signification statistique a été atteinte pour les valeurs $p < 0,05$. Toutes les analyses ont été effectuées à l'aide de Stata (version 17; StataCorp, College Station, TX, États-Unis).

II.3.3.3 Résultats

❖ *Tendances brutes de l'incidence mensuelle du paludisme simple et du paludisme grave*

Tableau 2, fig. Les figures 3 et 4 présentent respectivement les tendances brutes de l'incidence mensuelle du paludisme simple et du paludisme grave. Avant l'introduction de la CPS, l'incidence mensuelle brute du paludisme simple pendant la période de transmission élevée était plus élevée dans le groupe intervention et a diminué après les deux cycles de mise en œuvre de la CPS au niveau du groupe de comparaison (Tableau 2, Fig. 3). Avant la mise en place de la CPS, le groupe CPS présentait une forte évolution de l'incidence mensuelle brute du paludisme grave de la période de transmission élevée à celle du contrôle. Après le premier et le deuxième cycle de mise en œuvre du SMC, il y a eu un déclin dans le groupe SMC. (Tableau 2, Fig. 4).

Table 2 Number of uncomplicated and severe malaria cases, testing rate by group and year (August–November) among children under 5 years, 2013–2016

	2013	2014	2015	2016
Uncomplicated malaria cases				
Comparison group	209,816	216,047	244,386	487,104
SMC group	264,905	293,816	196,534	387,255
Severe malaria cases				
Comparison group	13,071	13,244	14,796	15,975
SMC group	14,887	17,483	12,746	14,360
Malaria suspected cases				
Comparison group	224,102	239,942	305,405	473,295
SMC group	279,499	311,299	209,756	364,534
Suspected malaria cases tested				
Comparison group	127,912	184,858	249,880	485,624
SMC group	166,575	259,008	219,637	351,972
Malaria confirmed cases				
Comparison group	115,702	168,142	221,919	394,583
SMC group	155,441	239,876	197,776	279,676
Testing rates (%)				
Comparison group	57.1	77.0	81.8	102.6
SMC group	59.6	83.2	104.7	96.6
Test positivity rate (%)				
Comparison group	90.5	91.0	88.8	81.3
SMC group	93.3	92.6	90.0	79.5

Trends of uncomplicated malaria incidence before and after SMC implementation

```
use smcstudy,clear
```

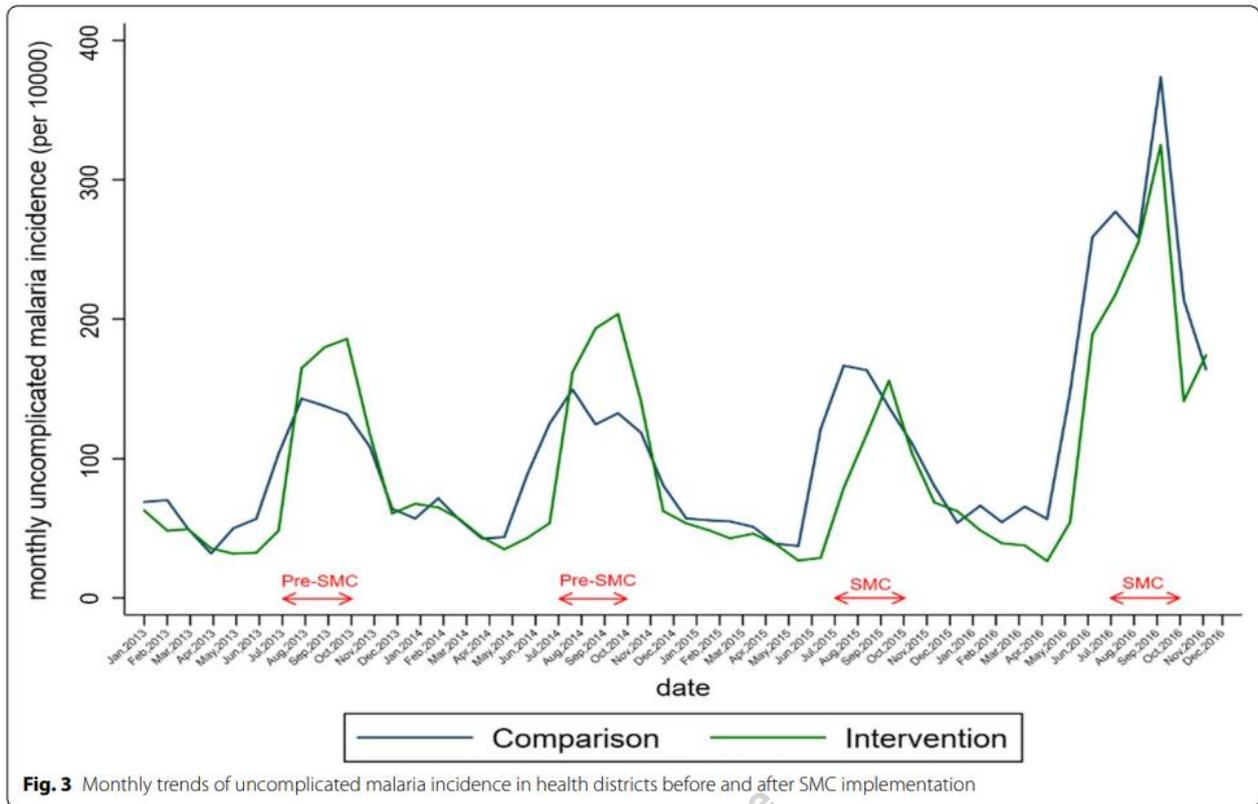
```
sort treat year month date
```

```
collapse umi,by(treat year month date)
```

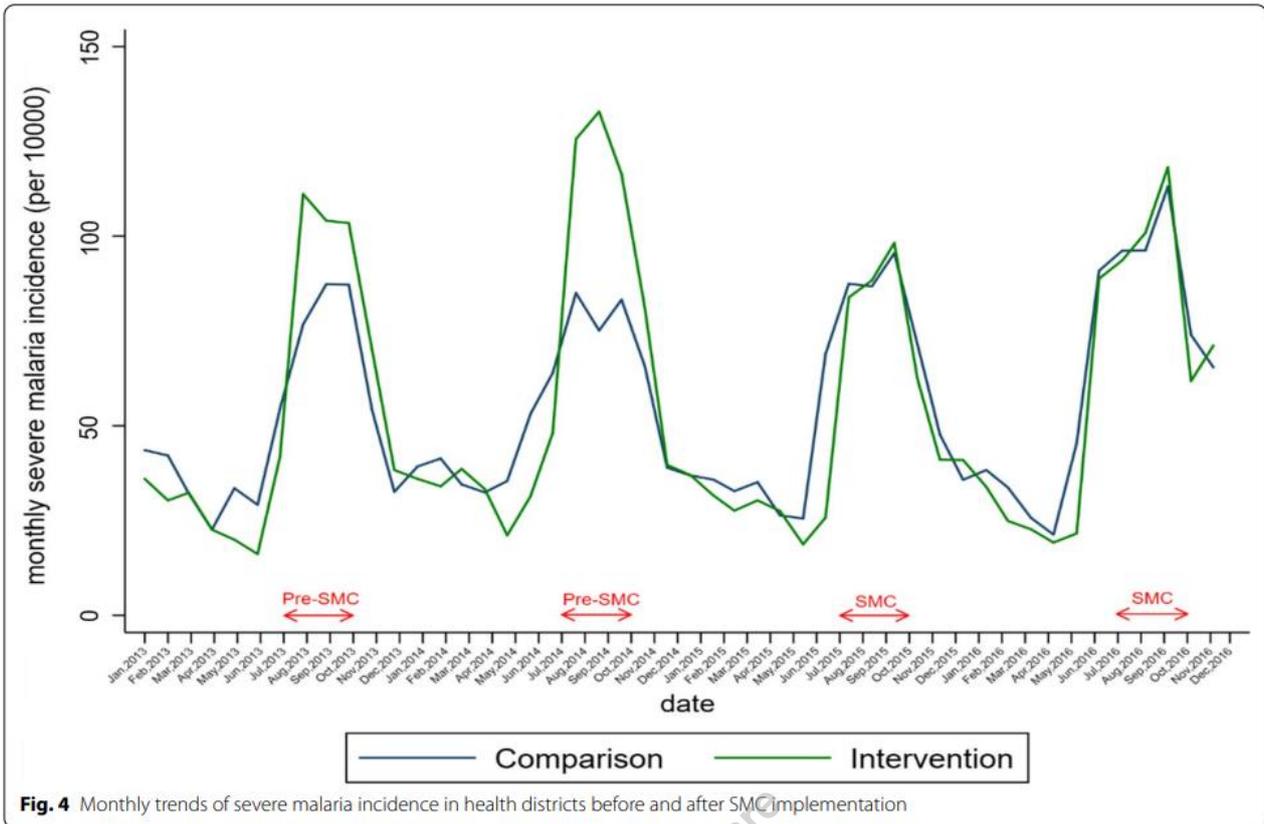
```
graph twoway line umi date if treat==0,lwidth(medium thick) tline(1Jun2015,lwidth(medium thick)) line umi date if treat==1,lwidth(medium thick) lcolor(green) ttext(60 1Jun2015 "", place(c)) ///
```

```
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white) margin(small)) ///
```

```
legend(lab(1 "Comparison") lab(2 "Intervention")) ytitle("(mean) monthly unc. malaria incidence")
```



```
//Trends of severe malaria incidence before and after SMC implementation
//*****
use smcstudy,clear
sort treat year month date
collapse smi,by(treat year month date)
graph twoway line smi date if treat==0,lwidth(medium thick) tline(1Jun2015,lwidth(medium thick))
line smi date if treat==1,lwidth(medium thick) lcolor(green) ttext(23 1Jun2015 "", place(c)) ///
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white) margin(small)) ///
legend(lab(1 "Comparison") lab(2 "Intervention")) ytitle("(mean) monthly severe malaria incidence")
```



```
//Trends of malaria mortality before and after SMC implementation
```

```
//*****
```

```
use smcstudy,clear
```

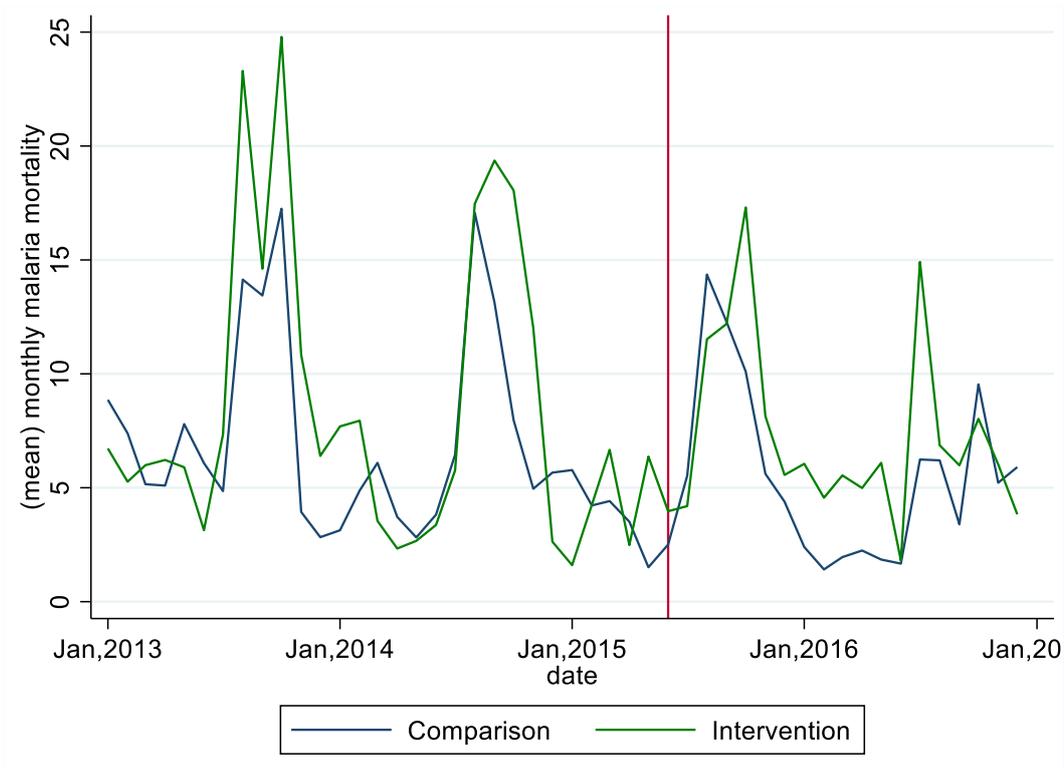
```
sort treat year month date
```

```
collapse mm,by(treat year month date)
```

```
graph twoway line mm date if treat==0,lwidth(medium thick) tline(1jun2015,lwidth(medium thick))|| line mm date if treat==1,lwidth(medium thick) lcolor(green) ttext(4.5 1jun2015 "", place(e)) ///
```

```
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white) margin(small)) ///
```

```
legend(lab(1 "Comparison") lab(2 "Intervention")) ytitle("(mean) monthly malaria mortality")
```



//PARALLEL TRENDS ASSUMPTION

//Trends of uncomplicated malaria incidence before and after SMC implementation

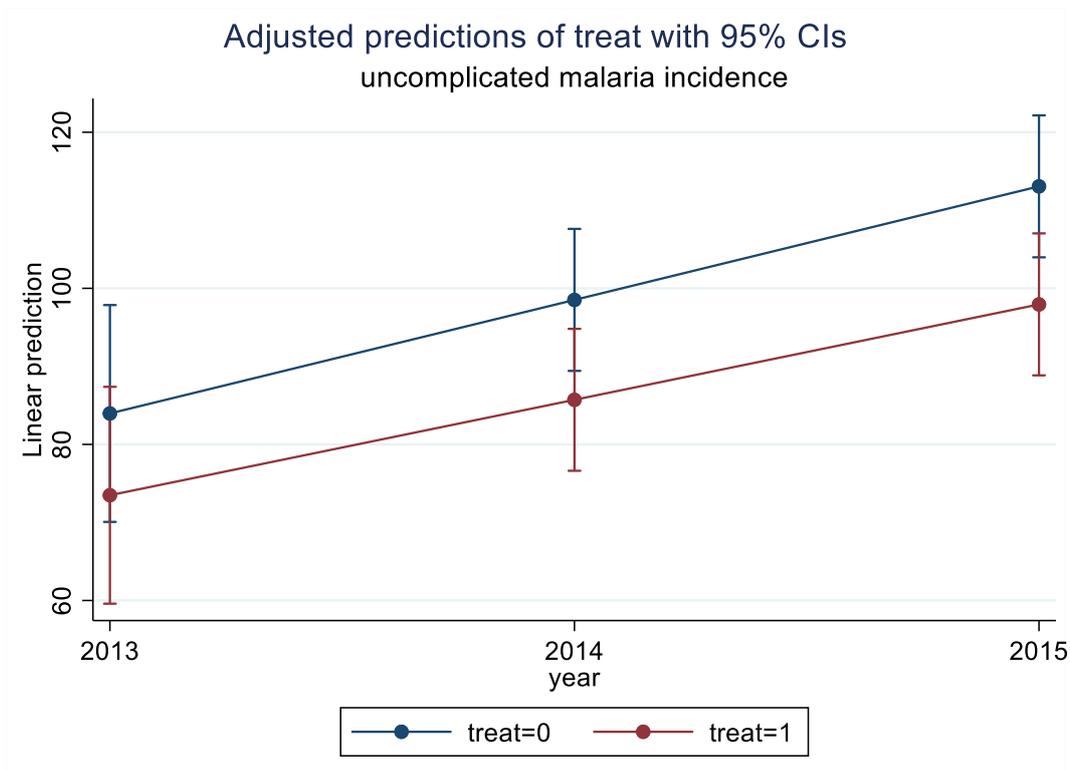
use smcstudy,clear

regress umi i.treat##c.year

margins treat, at(year = (2013(1)2015))

marginsplot,name(marginsplot, replace) saving(um1,replace) subtitle(uncomplicated malaria incidence) ///

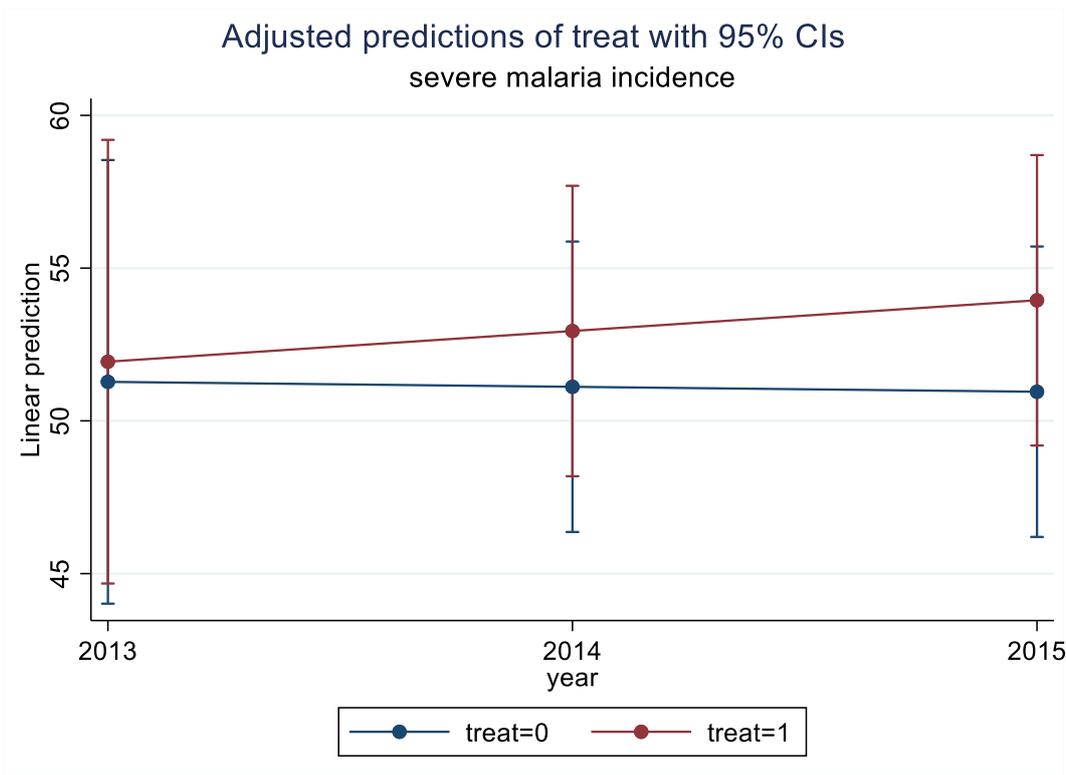
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white) margin(small))



```

*****
//Trends of severe malaria incidence before and after SMC implementation
*****
use smcstudy,clear
regress smi i.treat##c.year
margins treat, at(year = (2013(1)2015))
marginsplot,name(marginsplot, replace) saving(sm1,replace) subtitle(severe malaria
incidence) ///
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white) margin(small))

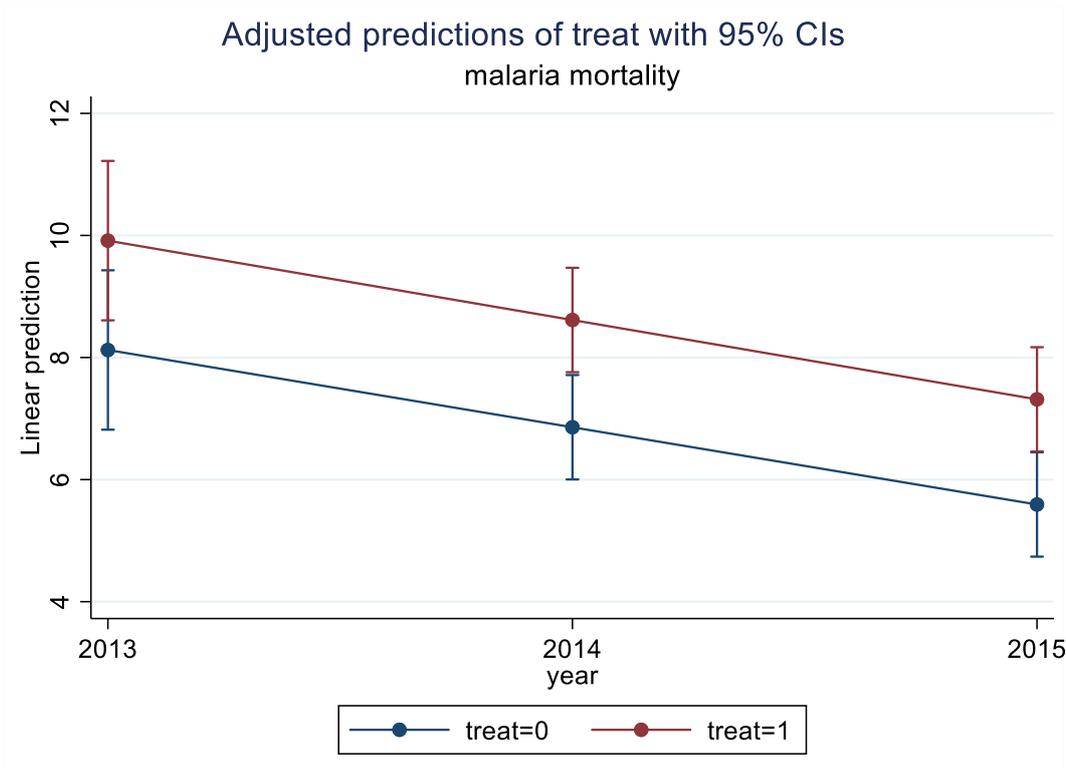
```



```

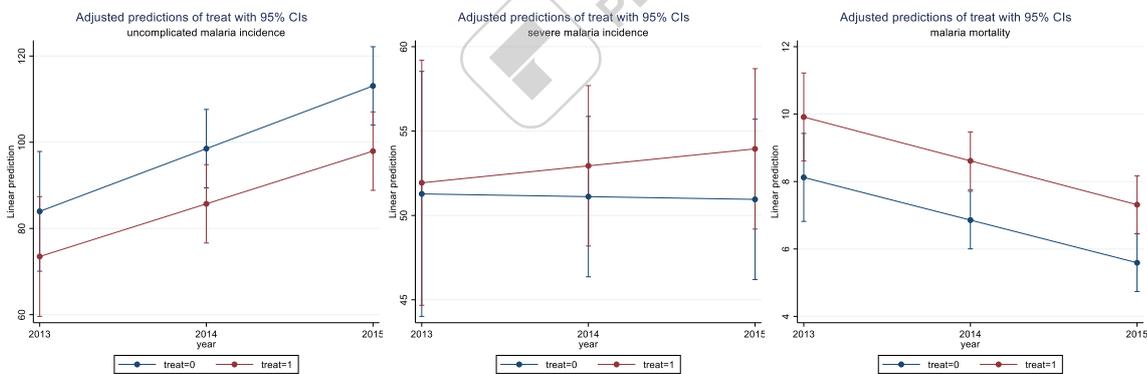
*****
//Trends of malaria mortality before and after SMC implementation
//*****
use smcstudy,clear
regress mm i.treat##c.year
margins treat, at(year = (2013(1)2015))
marginsplot,name(marginsplot, replace) saving(mm1,replace) subtitle(malaria mortality) ///
graphregion(fcolor(white) lcolor(white) ifcolor(white) ilcolor(white)
margin(small))

```



*****Tendance des trois outcome sur un même graphique

graph combine um1.gph sm1.gph mm1.gph, graphregion(fcolor(white) lcolor(white) /// ifcolor(white) ilcolor(white) margin(small)) rows(1) xsize(6) ysize(2)



❖ *Effet de la chimio prévention du paludisme saisonnier sur le fardeau du paludisme*
- *Paludisme simple*

gen logstdpop = log(u5)

*glm umal smc##time i.month1 i.year1 i.gratuite i.mal i.Distname, family(nb) link(log)
offset(logstdpop) eform vce(r)*

umal	IRR	std. err.	z	P>z	[95% conf.	interval]
1.smc	.8531814	.0813119	-1.67	0.096	.7078124	1.028406
1.time	1.729899	.1429213	6.63	0.000	1.471282	2.033975
smc#time						
1 1	.8280125	.0581152	-2.69	0.007	.7215961	.9501226
month1						
2	1.118438	.0590708	2.12	0.034	1.008452	1.240419
3	1.288228	.0655927	4.97	0.000	1.165875	1.423421
4	.7986981	.0424074	-4.23	0.000	.7197597	.8862938
year1						
2	1.124234	.0411467	3.20	0.001	1.046413	1.207843
3	.5927961	.0428753	-7.23	0.000	.5144465	.6830782
4	1	(omitted)				
1.gratuite	1	(omitted)				
1.mal	.9973939	.0699035	-0.04	0.970	.8693792	1.144259
Distname						
Boulsa	.6516117	.101875	-2.74	0.006	.4796334	.8852549
Diapaga	1.992981	.1583352	8.68	0.000	1.705604	2.328777
Gayeri	1.525183	.1210634	5.32	0.000	1.305439	1.781916
Kongoussi	1.402715	.1455341	3.26	0.001	1.144605	1.719028
Manga	1.529647	.1783199	3.65	0.000	1.217199	1.922299
Manni	1.701056	.1226392	7.37	0.000	1.476898	1.959236
Nanoro	1.486023	.1388507	4.24	0.000	1.237346	1.784679
Orodara	.7493558	.0902092	-2.40	0.017	.5918593	.9487628
Pama	1.681945	.1216063	7.19	0.000	1.459718	1.938003
Sapone	1.438554	.1781274	2.94	0.003	1.128567	1.833688
Sapouy	1.193637	.1145668	1.84	0.065	.9889471	1.440694
Sindou	2.359277	.2790162	7.26	0.000	1.871169	2.974712
Yako	1	(omitted)				
Ziniaré	1.743007	.1036331	9.34	0.000	1.551278	1.958432
Zorgho	2.082069	.1340203	11.39	0.000	1.835288	2.362033
_cons	.1074008	.0107112	-22.37	0.000	.0883316	.1305866
logstdpop	1	(offset)				

Paludisme grave

```
glm smal smc##time i.month1 i.year1 i.gratuite i.mal i.Distname, family(nb) link(log)
offset(logstdpop) eform vce(r)
```

smal	IRR	std. err.	z	P>z	[95% conf.	interval]
1.smc	1.046424	.1116922	0.43	0.671	.8488934	1.289919
1.time	.9846392	.0800324	-0.19	0.849	.8396353	1.154685
smc#time						
1 1	.94959	.0674585	-0.73	0.467	.8261656	1.091453
month1						
2	1.025031	.0529208	0.48	0.632	.9263831	1.134183
3	1.083407	.0568516	1.53	0.127	.9775187	1.200766
4	.6783686	.0367418	-7.16	0.000	.6100465	.7543424
year1						
2	1.149334	.0491659	3.25	0.001	1.056899	1.249852
3	1.042558	.0740725	0.59	0.557	.9070337	1.198332
4	1	(omitted)				
1.gratuite	1	(omitted)				
1.mal	.96609	.0728829	-0.46	0.647	.8333013	1.120039
Distname						
Boulsa	.822959	.1495143	-1.07	0.283	.5764118	1.174961
Diapaga	2.827149	.2242103	13.10	0.000	2.420155	3.302586
Gayeri	2.584172	.2141617	11.46	0.000	2.19674	3.039935
Kongoussi	2.248527	.2407317	7.57	0.000	1.822916	2.773508
Manga	1.400089	.1610781	2.93	0.003	1.117445	1.754224
Manni	2.098364	.1544121	10.07	0.000	1.816534	2.423919
Nanoro	2.250089	.2197431	8.30	0.000	1.85811	2.724757
Orodara	1.750342	.2162872	4.53	0.000	1.373856	2.229999
Pama	3.866175	.2936031	17.81	0.000	3.331501	4.486658
Sapone	1.91177	.2695813	4.60	0.000	1.450128	2.520372
Sapouy	1.989527	.228607	5.99	0.000	1.588336	2.492054
Sindou	2.933044	.3677126	8.58	0.000	2.294057	3.750013
Yako	1	(omitted)				
Ziniaré	1.446378	.0990471	5.39	0.000	1.264714	1.654138
Zorgho	2.095331	.1339901	11.57	0.000	1.848506	2.375114
_cons	.0045464	.0004981	-49.22	0.000	.0036678	.0056355
logstdpop	1	(offset)				

Mortalité

```
glm death smc##time i.month1 i.year1 i.gratuite i.mal i.Distname, family(nb) link(log)
offset(logstdpop) eform vce(r)
```

death	IRR	std. err.	z	P>z	[95% conf.	interval]
1.smc	5.387825	1.861369	4.87	0.000	2.737429	10.60435
1.time	.4039945	.0777282	-4.71	0.000	.2770799	.5890415
smc#time						
1 1	.7485457	.1300109	-1.67	0.095	.5325711	1.052105
month1						
2	.8579137	.0948046	-1.39	0.165	.6908452	1.065385
3	.9573323	.1178431	-0.35	0.723	.7521142	1.218545
4	.5053546	.0612227	-5.63	0.000	.3985428	.6407926
year1						
2	.9103951	.1042198	-0.82	0.412	.727422	1.139393
3	2.049988	.3349414	4.39	0.000	1.48825	2.823752
4	1	(omitted)				
1.gratuite	1	(omitted)				
1.mal	.7320595	.1369938	-1.67	0.096	.5072908	1.056418
Distname						
Boulsa	.7573195	.2137151	-0.99	0.325	.4355827	1.316703
Diapaga	1.850841	.3638165	3.13	0.002	1.25907	2.720748
Gayeri	.775585	.1884961	-1.05	0.296	.4816767	1.248829
Kongoussi	8.641248	2.371571	7.86	0.000	5.04622	14.79745
Manga	2.681267	.959431	2.76	0.006	1.32971	5.406589
Manni	.9367358	.237037	-0.26	0.796	.5704588	1.53819
Nanoro	4.535363	1.463892	4.68	0.000	2.409169	8.538012
Orodara	3.017678	.9635215	3.46	0.001	1.613953	5.642286
Pama	2.021608	.4868559	2.92	0.003	1.260972	3.24107
Sapone	2.957179	1.070868	2.99	0.003	1.454242	6.013381
Sapouy	2.719474	.8551177	3.18	0.001	1.468355	5.036615
Sindou	4.129537	1.740066	3.37	0.001	1.808131	9.431325
Yako	1	(omitted)				
Ziniaré	.4720103	.122865	-2.88	0.004	.2833886	.7861775
Zorgho	.8362753	.1403802	-1.07	0.287	.6018158	1.162077
_cons	.0000468	.0000158	-29.58	0.000	.0000242	.0000906
logstdpop	1	(offset)				

II.3.3.4 Discussion

Cette étude a évalué l'effet de la chimioprévention saisonnière du paludisme chez les enfants de moins de cinq ans sur la morbidité palustre à l'aide de données de routine au Burkina Faso de 2013 à 2016. Après deux cycles de CPS, nous avons observé une diminution significative de l'incidence des cas de paludisme simple et des cas de paludisme grave.

Ces résultats sont conformes à l'efficacité bien documentée de la CPS dans la réduction de l'incidence du paludisme grave.

Cette étude a également observé une augmentation de l'incidence du paludisme en 2016. Cela nécessite l'exploration d'une nouvelle stratégie comme une combinaison de CPS avec le vaccin antipaludique, le renforcement de la surveillance de la résistance à la SP, l'extension du groupe d'âge de la population cible, l'extension de la période à au moins 3 mois après la période de transmission élevée, la surveillance active au niveau communautaire.

Cette étude démontre la valeur ajoutée de l'utilisation de données de routine pour surveiller et évaluer efficacement l'effet sur les interventions antipaludiques au niveau infranational. Cela nécessite de continuer à renforcer les systèmes d'information de routine pour améliorer la qualité des données de surveillance du paludisme et l'utilisation de ces données pour éclairer la mise en œuvre du programme. Conformément à la High Burden High Impact Strategy (HBHI), une évaluation infranationale régulière de l'évolution du fardeau du paludisme est nécessaire pour définir un ensemble d'interventions appropriées ciblant une zone spécifique.



II.4. Séries temporelles interrompues

II.4.1 Introduction

L'analyse des séries chronologiques interrompues (STI) est une conception d'étude précieuse pour évaluer l'efficacité des interventions de santé au niveau de la population où elle a été mise en œuvre à un moment bien défini. Il est de plus en plus utilisé pour évaluer l'efficacité d'interventions en matière de santé publique. Bien que la conception partage de nombreuses propriétés des approches basées sur la régression dans d'autres études épidémiologiques, il existe une gamme de caractéristiques uniques des données de séries chronologiques qui nécessitent des considérations méthodologiques supplémentaires.

II.4.2 Le design des séries chronologiques interrompues

Une série chronologique est une séquence continue d'observations sur une population, prises à plusieurs reprises (normalement à intervalles égaux) au fil du temps. Dans une étude ITS, une série chronologique d'un résultat particulier d'intérêt est utilisée pour établir une tendance sous-jacente, qui est « interrompue » par une intervention à un moment connu. Le scénario hypothétique dans lequel l'intervention n'a pas eu lieu et la tendance se poursuit inchangée (c'est-à-dire : la tendance « attendue », en l'absence de l'intervention, compte tenu de la tendance préexistante) est appelé « contrefactuel ». Ce scénario contre-factuel fournit une comparaison pour l'évaluation de l'impact de l'intervention en examinant tout changement survenant dans la période post-intervention.

II.4.3 Étape 1 : un plan de série chronologique interrompu est-il approprié ?

La première décision lors de l'examen d'un STI est de savoir s'il s'agit d'une conception appropriée pour l'évaluation particulière en question. Cela dépend de la nature de l'intervention et du résultat de l'intérêt, ainsi que du type de données disponibles :

Les STI d'intervention exigent une différenciation claire de la période précédant l'intervention et de la période postérieure à l'intervention. Dans certaines évaluations, il peut être difficile de définir quand l'intervention a commencé et de différencier les effets des différentes composantes. Cela ne nécessite pas nécessairement que l'intervention soit introduite du jour au lendemain, mais la période de mise en œuvre doit être bien définie afin qu'elle puisse être considérée séparément.

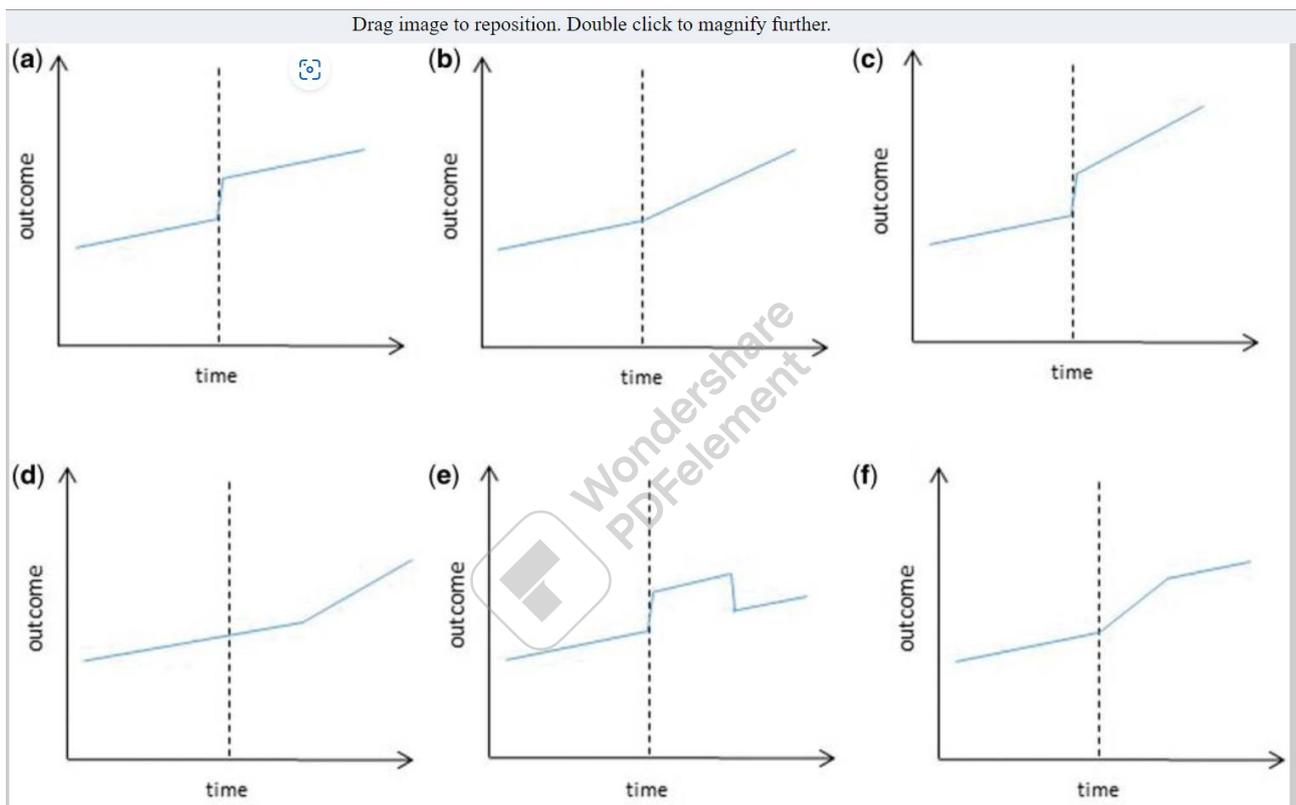
La mise en œuvre de l'exemple d'intervention a été très claire avec l'administration de la CPS en juin 2015 à tous les enfants de 3-59 mois dans les districts d'interventions retenus.

Les résultats peuvent prendre diverses formes telles que des comptes, des données continues ou des variables binaires. Les STI fonctionnent mieux avec des résultats à court terme qui devraient changer soit relativement rapidement après la mise en œuvre d'une intervention, soit après un décalage clairement défini.

II.4.4 Étape 2 : proposer le modèle d'impact

Une fois qu'une conception de STI est choisie, l'étape suivante consiste à émettre l'hypothèse de l'impact de l'intervention sur le résultat si elle était efficace, en particulier si le changement était un changement progressif du gradient de la tendance, un changement dans le niveau ou

les deux, et si le changement suivrait immédiatement l'intervention ou s'il y aura une période de décalage avant qu'un effet ne soit attendu. Des exemples de certains modèles d'impact possibles sont illustrés dans le graphique ci-dessous. Il est important que cette décision soit prise a priori sur la base de la littérature existante et des connaissances de l'intervention et du mécanisme par lequel elle est censée agir sur le résultat. Lorsque les connaissances existantes sur l'intervention sont limitées, le choix du modèle d'impact le plus approprié peut être difficile et peut nécessiter une analyse exploratoire de données alternatives. Il est déconseillé de s'appuyer sur les données ou sur les résultats pour sélectionner le meilleur modèle d'impact, car cela augmente la probabilité qu'un effet soit détecté en raison de fluctuations aléatoires ou du hasard, et les conclusions artéfactuelles qui en découlent sur l'effet de l'intervention.



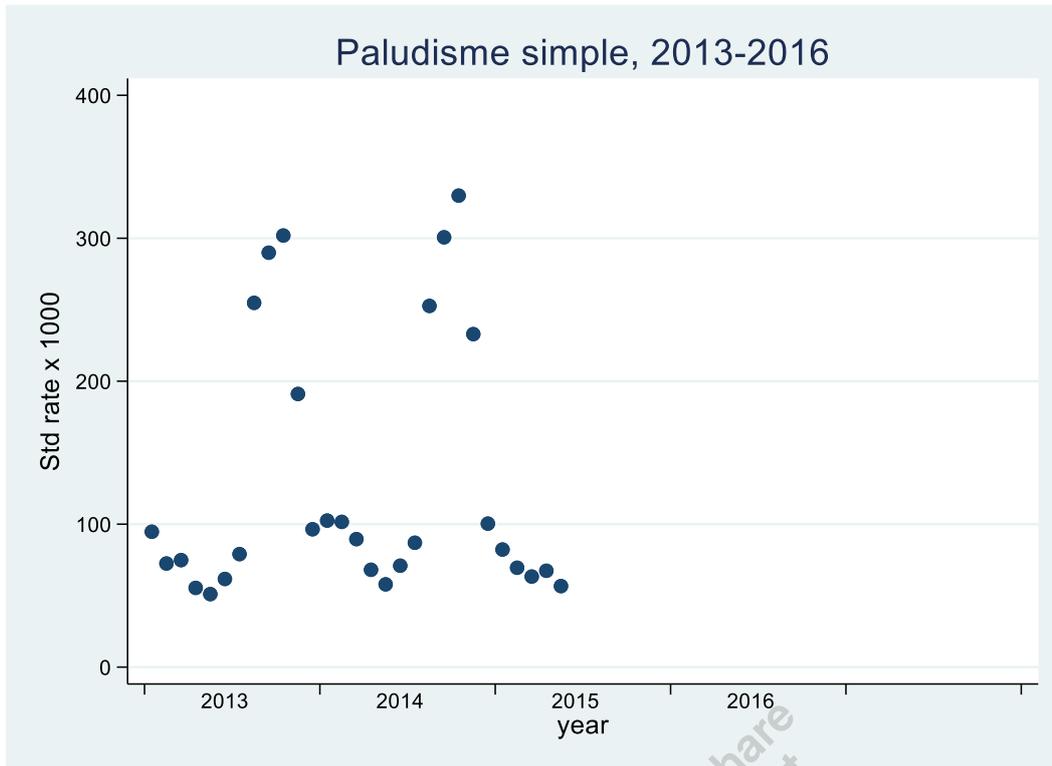
Exemples de modèles d'impact utilisés dans les STI

a) Changement de niveau ; b) Changement de pente; c) Changement de niveau et de pente; d) Changement de pente à la suite d'un décalage; e) Changement temporaire de niveau; f) Changement temporaire de pente entraînant un changement de niveau.

II.4.5 Étape 3 : analyse descriptive

Comme pour toutes les analyses statistiques, des statistiques sommaires initiales et des graphiques devraient être entrepris pour familiariser les chercheurs avec les données. Cela devrait inclure un nuage de points de la série chronologique, comme indiqué dans le **graphique** ci-dessous, ce qui peut aider à identifier la tendance sous-jacente, les tendances saisonnières et les valeurs aberrantes. Des analyses descriptives plus traditionnelles, telles que des résumés et

des comparaisons bivariées entre le résultat et les facteurs de confusion potentiels variant dans le temps, ainsi que de simples comparaisons avant-après, sont recommandées.



II.4.6 Étape 4 : analyse de régression

Au moins trois variables sont requises pour une analyse des STI :

- i. T : le temps écoulé depuis le début de l'étude, l'unité représentant la fréquence à laquelle les observations sont effectuées (par exemple, mois ou année);
- ii. X_t : une variable fictive indiquant la période pré-intervention (codée 0) ou la période post-intervention (codée 1) ;
- iii. Y_t : le résultat à l'instant t .

Dans les analyses ITS standard, le modèle de régression segmenté suivant est utilisé :

$$Y_t = \beta_0 + \beta_1 T + \beta_2 X_t + \beta_3 T X_t$$

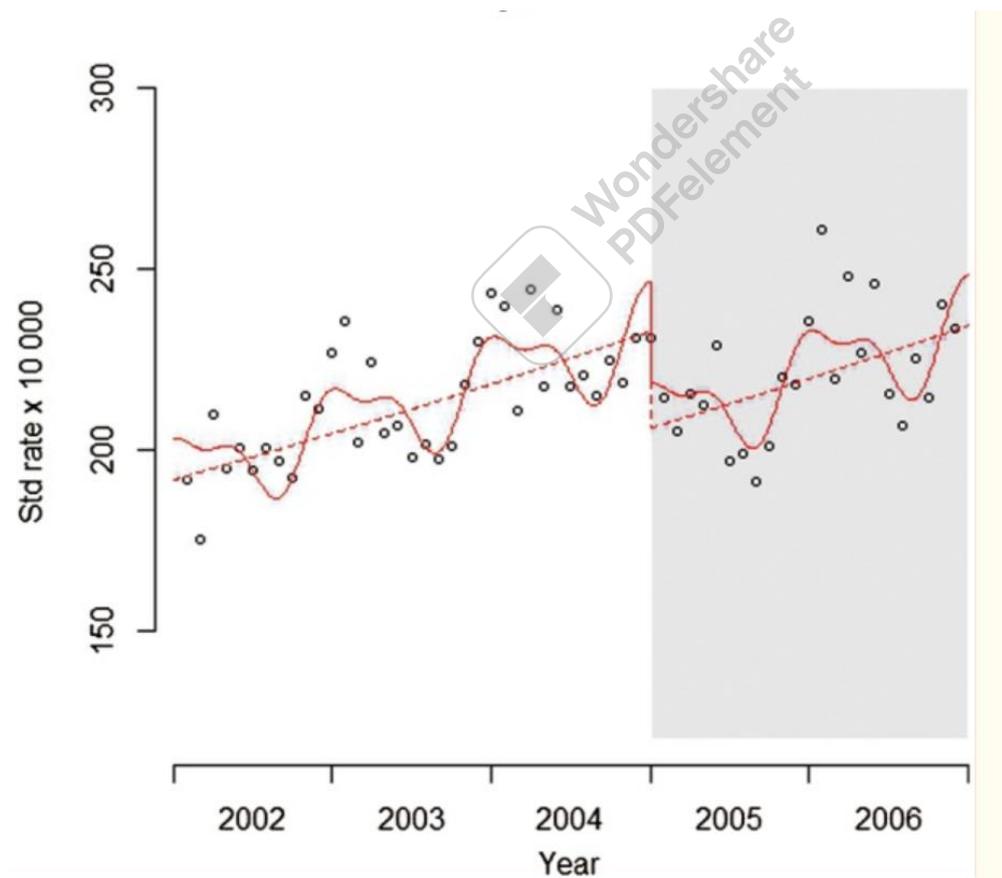
où β_0 représente le niveau de référence à $T = 0$, β_1 est interprété comme le changement de résultat associé à une augmentation de l'unité de temps (représentant la tendance sous-jacente avant l'intervention), β_2 est le changement de niveau suivant l'intervention et β_3 indique le changement de pente suivant l'intervention (en utilisant l'interaction entre le temps et l'intervention : TX_t). Le modèle de régression ci-dessus représente le modèle d'impact (c); les modèles (a) et (b) peuvent facilement être spécifiés en excluant les termes $\beta_3 TX_t$ ou $\beta_2 X_t$ respectivement. Les modèles d'impact (d)-(f) nécessitent des spécifications variables légèrement plus complexes.

II.4.6 Étape 5 : résoudre les problèmes méthodologiques

Alors que le modèle de base mis en œuvre jusqu'à présent fournit une indication de l'association potentielle entre l'intervention et le résultat, il existe un certain nombre de problèmes distincts avec les données de séries chronologiques qui peuvent devoir être abordés afin d'améliorer la robustesse de l'analyse.

Saisonnalité : De nombreuses maladies et autres critères de jugement ont une tendance saisonnière, ce qui est évident dans les données de paludisme. La saisonnalité peut causer deux problèmes :

- ❖ Premièrement, s'il y a une répartition inégale des mois avant et après l'intervention, comme une proportion plus élevée de mois d'hiver, cela pourrait biaiser les résultats, en particulier dans l'analyse des séries courtes.
- ❖ Deuxièmement, les résultats d'un mois ont tendance à être plus semblables à ceux des mois voisins à la même période de l'année, ce qui entraîne une autocorrélation et une surdispersion. Il existe une gamme de méthodes pour contrôler la saisonnalité et d'autres tendances à long terme ; il s'agit notamment d'un modèle en utilisant des fonctions plus complexes telles que les termes de Fourier (paires de fonctions sinus et cosinus).



Avantages : L'un des points forts des études sur les STI est qu'elles ne sont généralement pas affectées par des variables confondantes typiques qui restent assez constantes, telles que la répartition par âge de la population ou le statut socioéconomique, car celles-ci ne changent que relativement lentement au fil du temps et sont normalement prises en compte lors de la

modélisation de la tendance sous-jacente à long terme. Néanmoins, les STI peuvent être affectés par des facteurs de confusion variant dans le temps qui changent plus rapidement. La saisonnalité peut être considérée comme un facteur de confusion variant dans le temps ; d'autres peuvent inclure les niveaux d'une maladie infectieuse particulière qui est sujette à des éclosions, des événements météorologiques, etc. Lorsque de tels facteurs de confusion variant dans le temps ont été mesurés, ils peuvent être contrôlés en incluant des variables les représentant dans le modèle de régression, comme c'est généralement le cas dans d'autres analyses épidémiologiques. Une catégorie particulière de facteurs de confusion variant dans le temps sont d'autres événements qui se produisent à peu près au même moment que l'intervention et qui peuvent influencer le résultat. Ceux-ci peuvent inclure d'autres interventions simultanées ciblant le même résultat, ou des facteurs de risque pour ce résultat, ou des événements naturels qui pourraient affecter le résultat.

Sur-dispersion Une hypothèse de la distribution de Poisson est que la variance est égale au nombre attendu. Cependant, dans les analyses de données réelles, la variance a souvent tendance à être plus grande (un phénomène connu sous le nom de sur dispersion), ce qui conduirait à une **estimation incorrecte des erreurs-types**. Un ajustement de mise à l'échelle est donc effectué pour corriger le modèle afin de corriger cela.

Autocorrélation Une deuxième hypothèse des modèles de régression standard est que les observations sont indépendantes. Cette hypothèse est souvent violée dans les données de séries chronologiques parce que les observations consécutives ont tendance à être plus similaires les unes aux autres que celles qui sont plus éloignées, un phénomène connu sous le nom **d'autocorrélation**. Heureusement, dans de nombreuses données épidémiologiques, l'autocorrélation s'explique en grande partie par d'autres variables, en particulier la saisonnalité (discutée ci-dessus) ; par conséquent, après avoir contrôlé ces facteurs, l'autocorrélation résiduelle est rarement un problème. Néanmoins, l'autocorrélation devrait toujours être évaluée en examinant le diagramme des résidus et la fonction d'autocorrélation partielle et, lorsque les données sont normalement distribuées, en effectuant des tests tels que le test de Breusch-Godfrey. Lorsque l'autocorrélation résiduelle demeure, elle devrait être ajustée pour tenir compte de méthodes telles que la régression de Prais ou la moyenne mobile intégrée autorégressive (ARIMA).

II.4.7 Étape 6 : vérification du modèle et analyses de sensibilité

Une gamme de techniques de vérification des modèles a été décrite ci-dessus, y compris le traçage des résidus et des fonctions d'autocorrélation partielle. En outre, des analyses de sensibilité peuvent être effectuées pour tester l'impact de la variation d'une gamme d'hypothèses du modèle, telles que différents décalages, types de modèles d'impact ou approches d'ajustement pour la saisonnalité.

II.4.8 Travaux pratiques

Titre : Évaluation de l'effet de la chimio prévention du paludisme saisonnier sur le fardeau du paludisme chez les enfants de moins de 5 ans au Burkina Faso

II.4.8.1 Introduction

Idem pour TP1

II.4.8.2 Méthodes

❖ *Conception de l'étude*

- Période retenue pour l'étude 2013-2016
- Période de mise en œuvre de la CPS : juin 2015
- Pas de groupe control

❖ *Sélection des districts de santé*

- Districts n'ayant pas bénéficié de la CPS avant juin 2015 et qui a bénéficié à partir de juin 2015

❖ *Sources des données*

- Données sur les cas de paludisme : Entrepôt de données sanitaire du ministère de la santé
- Données de population des enfants de moins de 5 ans : projection démographique de l'INSD
- Covariables : Gratuité des soins de santé (0 avant 2016 et 1 après 2016)

❖ *L'analyse des données*

- Descriptives
- Modélisation

II.4.8.3 Résultats

❖ *Analyse descriptive*

L'examen des données est une première étape importante. L'examen de la tendance **pré-intervention** peut donner une indication de la stabilité de la tendance dans le temps, si un modèle linéaire est susceptible d'être approprié et s'il semble y avoir une tendance saisonnière.

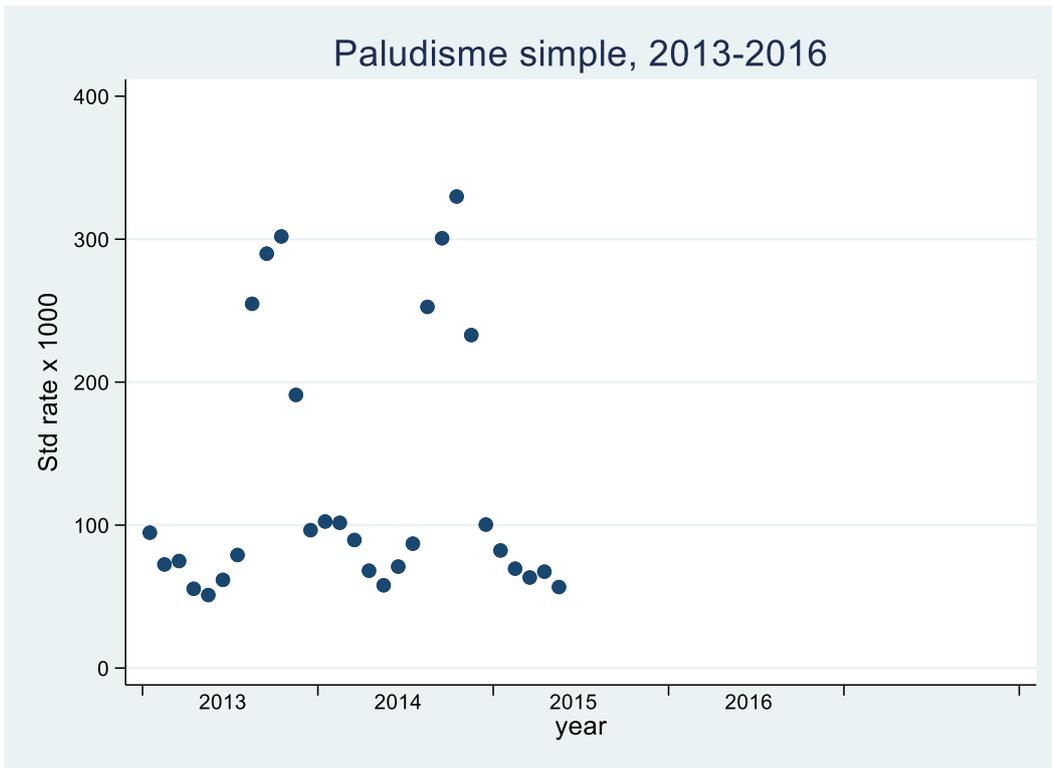
Etape 1:

```
gen rate = umal/pop*10^3
```

```
twoway (scatter rate time) if smc==0, title("Paludisme simple, 2013-2016") ytitle(Std rate x 1000)
yscale(range(0 .)) ylabel(#5, labsize(small) angle(horizontal)) ///
```

```
xtick(0.5(12)60.5) xlabel(6"2013" 18"2014" 30"2015" 42"2016", noticks labsize(small))
xtitle(year)
```

```
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques
publiques\Regression_seg\Graphd1.tif", as(tif) replace
```



❖ **Modèle de régression de Poisson**

Nous utilisons également un modèle de Poisson, car nous utilisons des données de comptage pour ce faire, nous modélisons directement les données de comptage. Nous utilisons ensuite le log de la population (log transformée) comme variable de décalage afin de la reconvertir en taux.

*log transformer la population

Etape 2:

gen logstdpop = log(pop)

glm umal smc time, family(poisson) link(log) offset(logstdpop) eform

umal	IRR	std. err.	z	P>z	[95% conf. interval]
smc	.7628125	.0021091	-97.92	0.000	.7586899 .7669575
time	1.026619	.0001064	253.52	0.000	1.026411 1.026828
_cons	.0846268	.0001718	1216.28	0.000	.0842907 .0849642
logstdpop	1	(offset)			

Nous générons des valeurs prédites basées sur le modèle afin de créer un tracé du modèle :

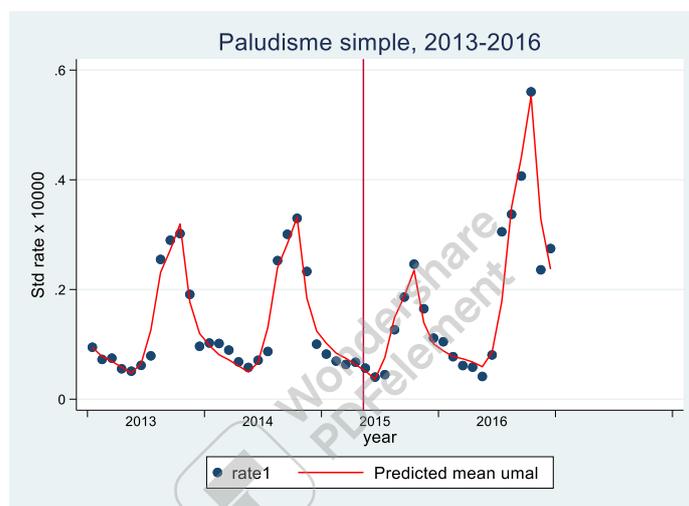
Les valeurs prédites peuvent ensuite être tracé avec un nuage de points pour vérifier la qualité de la prédiction du modèle.

predict pred, nooffset

* Cela peut ensuite être tracé avec un nuage de points :

Etape 3:

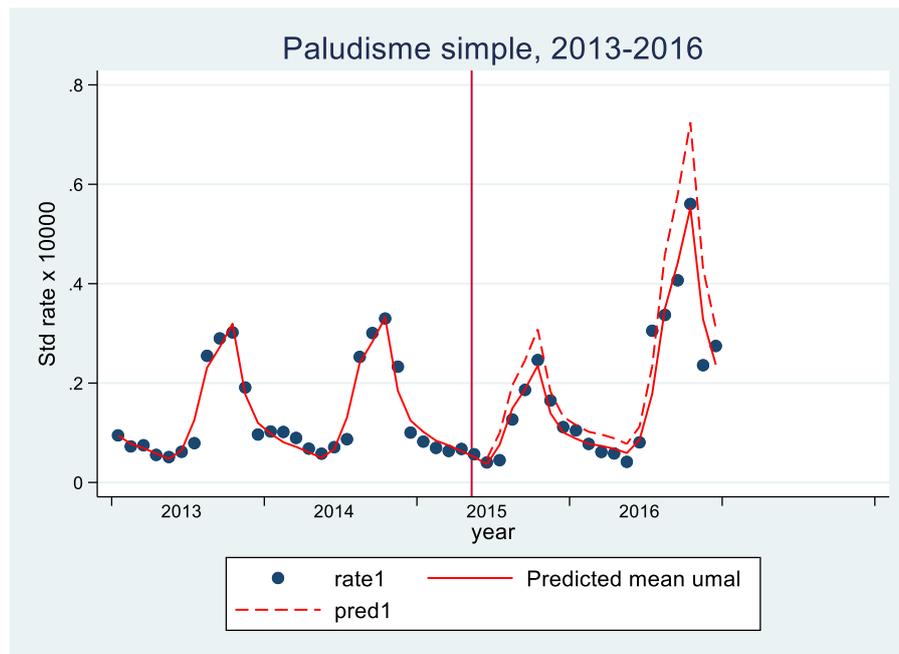
```
gen rate1 = umal/pop /*to put rate in same scale as count in model */
twoway (scatter rate1 time) (line pred time, lcolor(red)) , title("Paludisme simple, 2013-2016")
///
ytitle(Std rate x 10000) yscale(range(0.)) ylabel(#5, labsize(small) angle(horizontal)) ///
xtick(0.5(12)60.5) xlabel(6"2013" 18"2014" 30"2015" 42"2016", noticks labsize(small))
xtitle(year) ///
xline(28.8)
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques
publiques\Regression_seg\Graphd2.tif", as(tif) replace
```



Générer le contre-factuel en supprimant l'effet de l'intervention ($_b[\text{smc}]$) pour la période post-intervention et ajouter le contre-factuel au tracer.

Etape 4:

```
gen pred1 = pred/exp[_b[smc]] if smc==1
twoway (scatter rate1 time) (line pred time, lcolor(red)) (line pred1 time, lcolor(red)
lpattern(dash)), title("Paludisme simple, 2013-2016") ///
ytitle(Std rate x 10000) yscale(range(0.)) ylabel(#5, labsize(small) angle(horizontal)) ///
xtick(0.5(12)60.5) xlabel(6"2013" 18"2014" 30"2015" 42"2016", noticks labsize(small))
xtitle(year) ///
xline(28.8)
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques
publiques\Regression_seg\Graphd3.tif", as(tif) replace
```



Étape 5: problèmes méthodologiques

(a) Permettre la surdispersion

Dans le modèle ci-dessus, nous n'avons pas autorisé la surdispersion pour ce faire, nous pouvons ajouter le paramètre d'échelle ($x2$) dans le modèle qui permet à la variance d'être proportionnelle plutôt qu'égal à la moyenne.

Étape 5:

```
glm umal smc time, family(poisson) link(log) offset(logstdpop) scale(x2) eform
```

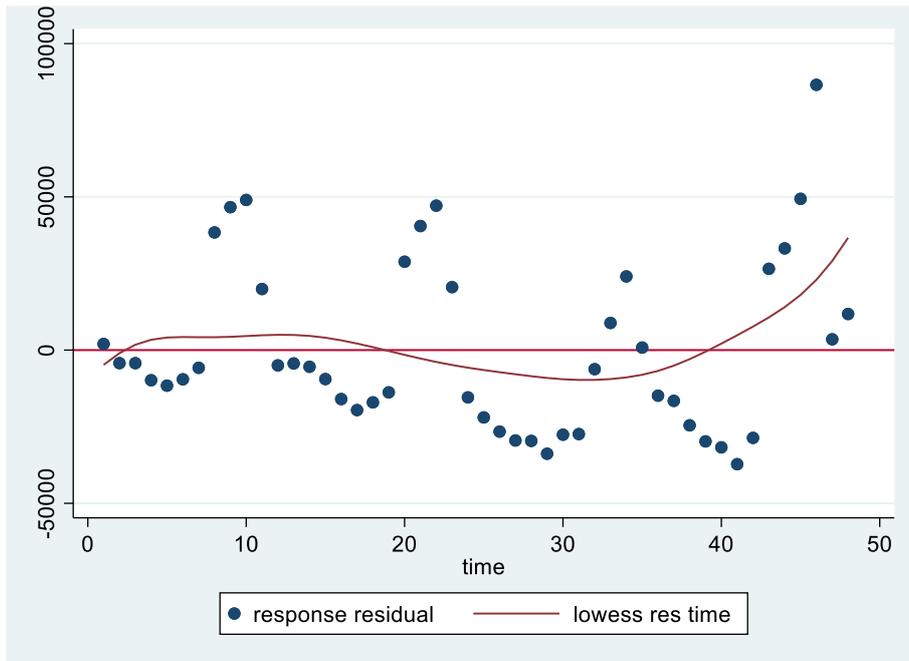
Étape 6

(b) Vérification du modèle et autocorrélation

```
predict res, r
```

```
twoway (scatter res time)(lowess res time),yline(0)
```

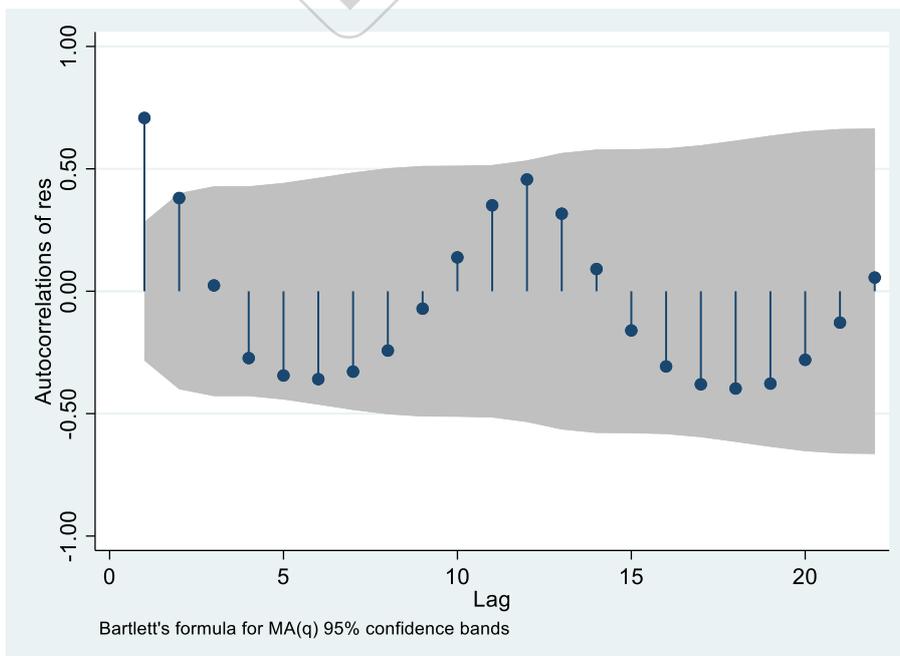
```
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques publiques\Regression_seg\Graphd4.tif", as(tif) replace
```



Vérifiez davantage l'autocorrélation en examinant les fonctions d'autocorrélation et d'autocorrélation partielle

Etape 6 suite

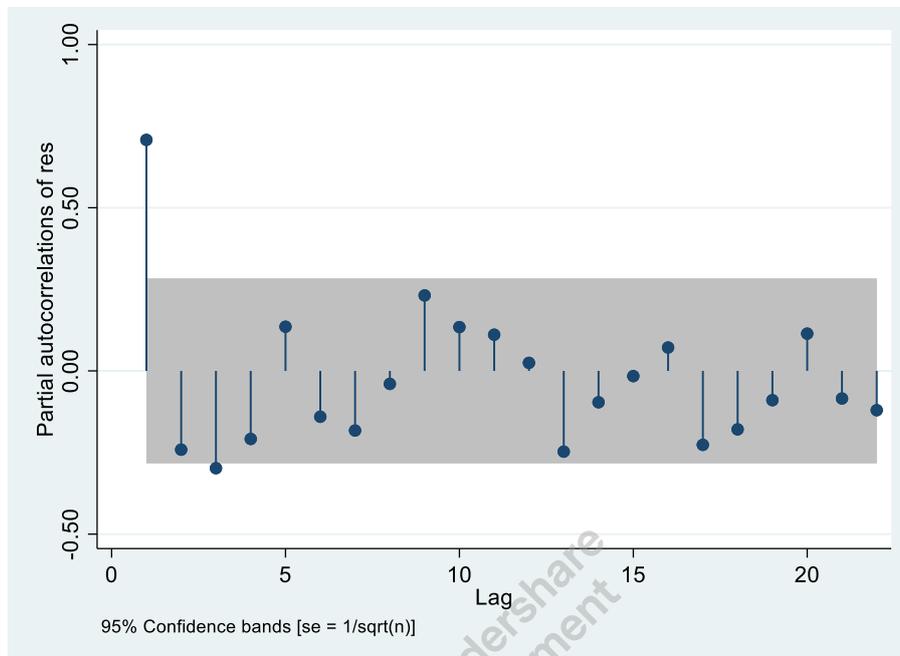
```
tsset time
ac res
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques publiques\Regression_seg\Graphd5.tif", as(tif) replace
```



Autocorrélation partielle

pac res, yw

graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques publiques\Regression_seg\Graphd6.tif", as(tif) replace



❖ **Paludisme est saisonnier**

(c) ajuster pour la saisonnalité

- Installation du paquet "circulaire". Pour rechercher des packages, sélectionnez Aide > SJ et Programmes écrits par l'utilisateur et cliquez sur rechercher */
- Nous devons créer une variable de degrés pour le temps divisé par le nombre de points dans le temps dans une année (c'est-à-dire 12 pour les mois)

Etape 7:

```
gen degrees=(time/12)*360
```

```
fourier degrees, n(2)
```

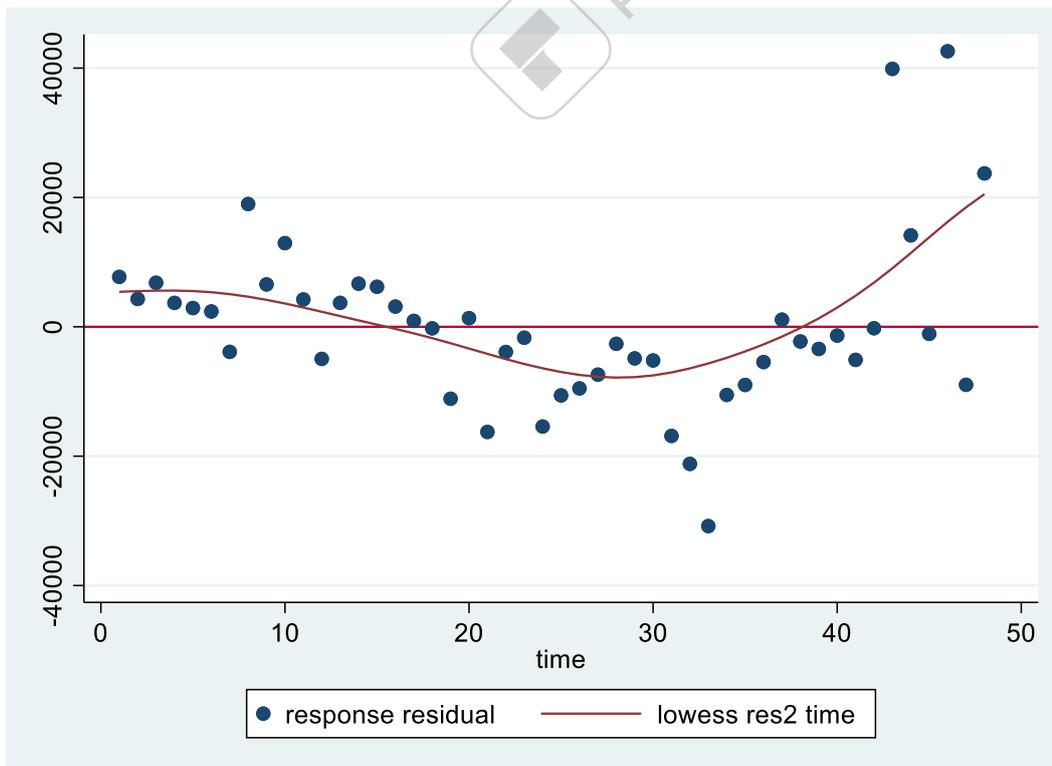
```
glm umal smc cos* sin* time, family(poisson) link(log) offset(logstdpop) scale(x2) eform
```

umal	IRR	std. err.	z	P>z	[95% conf. interval]
smc	.6031086	.1162746	-2.62	0.009	.4133239 .8800361
cos_1	1.374348	.1129391	3.87	0.000	1.169898 1.614528
cos_2	.7988917	.0585947	-3.06	0.002	.691921 .9224
sin_1	.462024	.0365255	-9.77	0.000	.3957059 .5394566
sin_2	1.023299	.0756085	0.31	0.755	.885339 1.182756
time	1.026681	.0073666	3.67	0.000	1.012343 1.041221

Vérification de 'autocorrélation

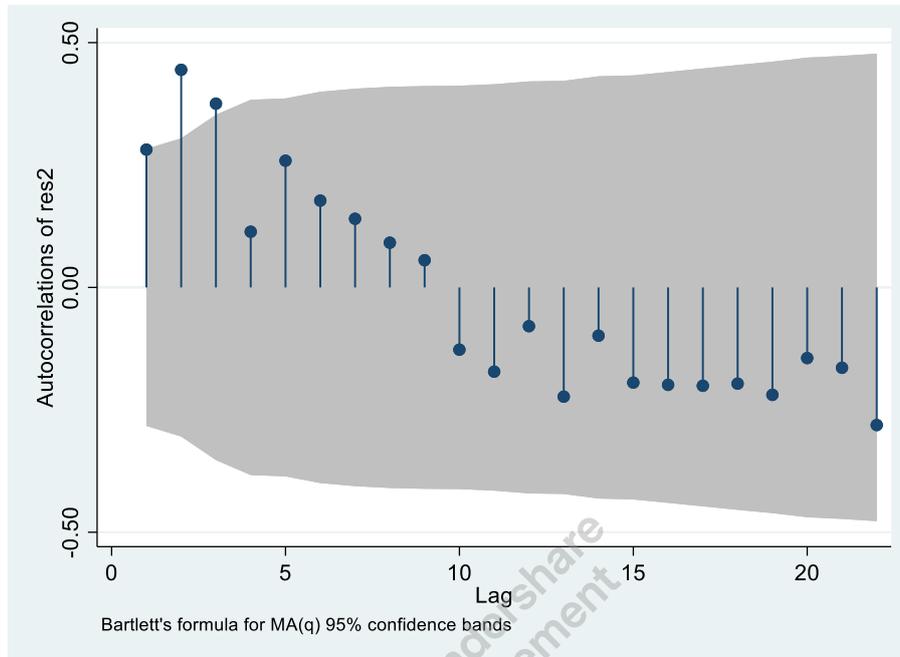
```
predict res2, r
```

```
twoway (scatter res2 time)(lowess res2 time),yline(0)
```

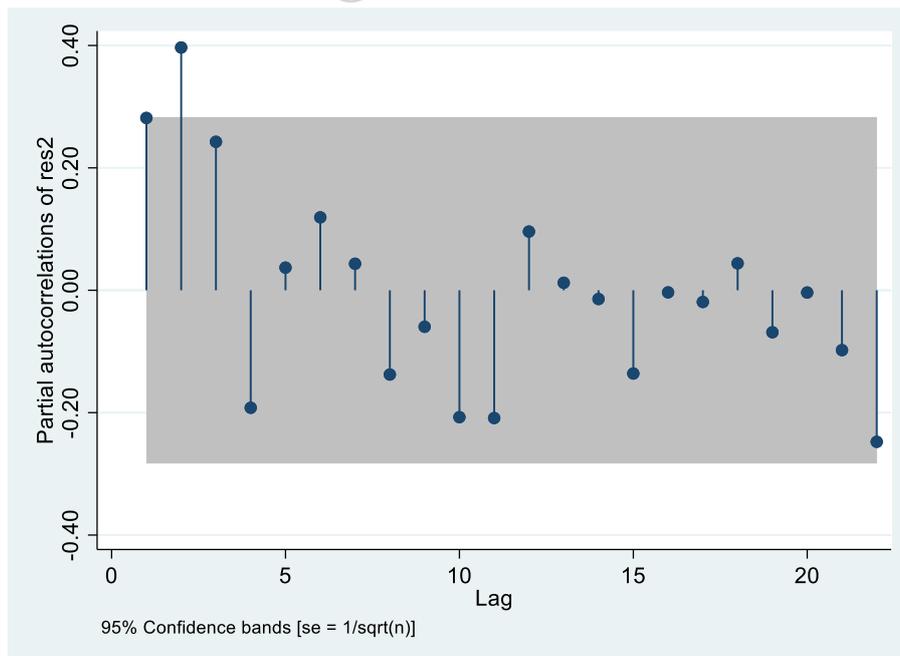


Vérification de 'autocorrélation

```
tsset time
ac res2
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques publiques\Regression_seg\Graphd7.tif", as(tif) replace
```



Autocorrélation partielle
pac res2, yw
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques publiques\Regression_seg\Graphd8.tif", as(tif) replace

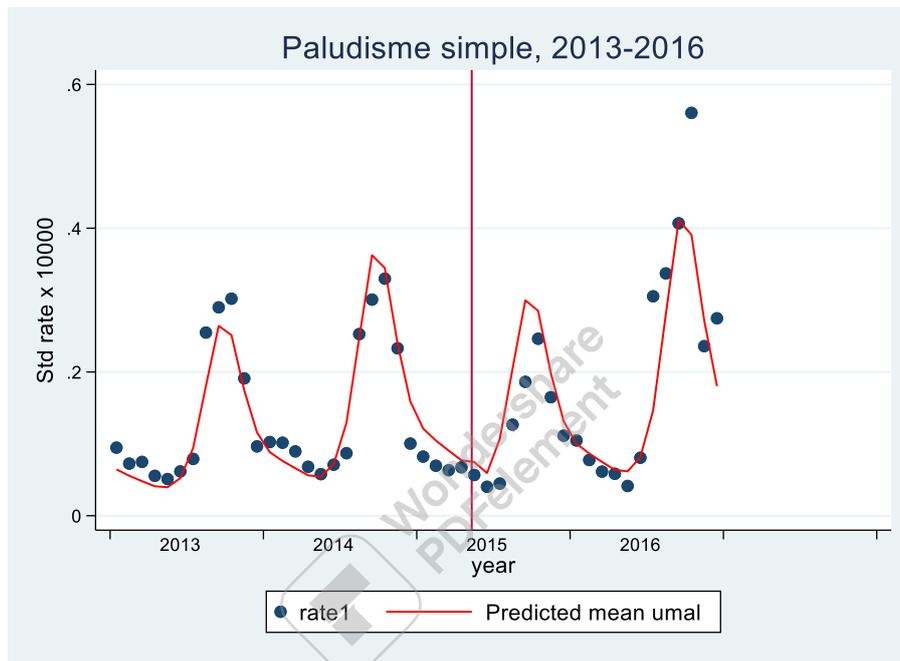


❖ **Prédire et tracer le modèle désaisonnalisé**

```

predict pred2, nooffset
twoway (scatter rate1 time) (line pred2 time, lcolor(red)), title("Paludisme simple, 2013-2016")
///
yttitle(Std rate x 10000) yscale(range(0.)) ylabel(#5, labsize(small) angle(horizontal)) ///
xtick(0.5(12)60.5) xlabel(6"2013" 18"2014" 30"2015" 42"2016", noticks labsize(small))
xtitle(year) ///
xline(28.8)
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques
publiques\Regression_seg\Graphd9.tif", as(tif) replace

```



Il est parfois difficile de voir clairement le changement graphiquement dans le modèle désaisonnalisé, il peut donc être utile de tracer une ligne droite comme si tous les mois étaient la moyenne pour produire une tendance « désaisonnalisée ».

```

egen avg_cos_1 = mean(cos_1)
egen avg_sin_1 = mean(sin_1)
egen avg_cos_2 = mean(cos_2)
egen avg_sin_2 = mean(sin_2)
drop cos* sin*
rename avg_cos_1 cos_1
rename avg_sin_1 sin_1
rename avg_cos_2 cos_2
rename avg_sin_2 sin_2
** cela peut ensuite être ajouté au tracé sous forme de ligne pointillée

```

```
predict pred3, noffset
```

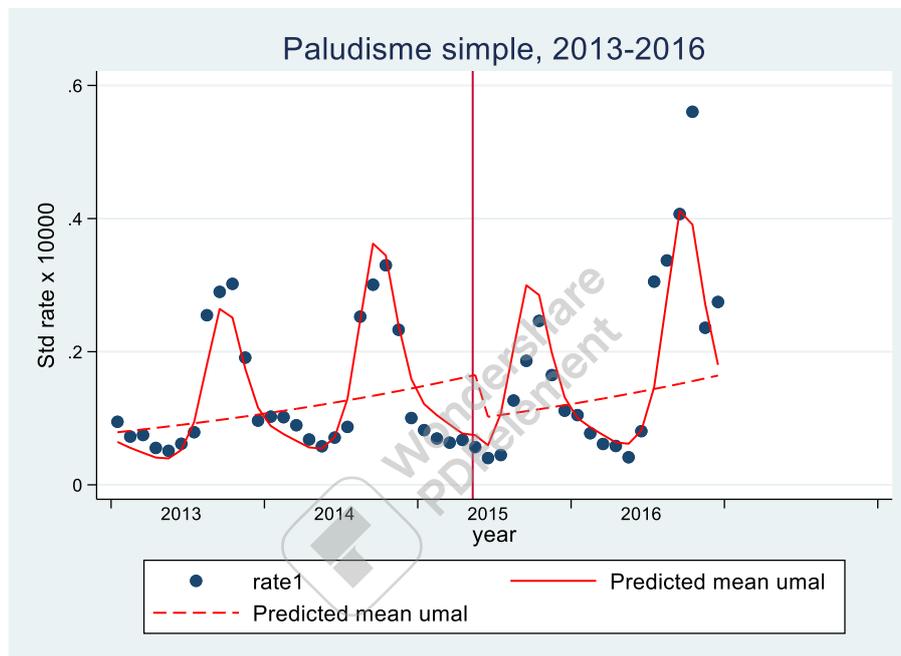
```
twoway (scatter rate1 time) (line pred2 time, lcolor(red)) (line pred3 time, lcolor(red)
lpattern(dash)), title("Paludisme simple, 2013-2016") ///
```

```
ytitle(Std rate x 10000) yscale(range(0.)) ylabel(#5, labsize(small) angle(horizontal)) ///
```

```
xtick(0.5(12)60.5) xlabel(6"2013" 18"2014" 30"2015" 42"2016", noticks labsize(small))
xtitle(year) ///
```

```
xline(28.8)
```

```
graph export "D:\Expertise France_Bonkongou\Formation evaluation des politiques
publiques\Regression_seg\Graphd10.tif", as(tif) replace
```



Annexes

Annexe1 : modules de formation

Annexe2 : Chronogramme

Annexe3 : Liste de présence

Annexe4 : Bases de données

