

Master Économétrie et Statistiques, parcours Économétrie Appliquée

Mémoire de Master 1

# L'immigration est-elle un facteur d'augmentation de la criminalité ? Une analyse en panel des départements français

MARTERET Achille

Sous la direction de Pistolesi Nicolas

Année académique : 2024-2025

## **Remerciements**

*Je suis reconnaissant envers ma famille pour son soutien tout au long de la réalisation de ce mémoire.*

*Je tiens à exprimer ma gratitude à mes amies, qui ont facilité les échanges autour de mon sujet, ont apporté des idées et suggestions d'améliorations utiles pour l'enrichissement mon travail.*

*Je remercie Monsieur Nicolas Pistolesi pour m'avoir proposé ce sujet de mémoire et pour m'avoir accompagné dans son développement.*

# **Analyse de l'effet de l'immigration sur la criminalité et délinquances des départements français entre 2006 et 2021, à travers une modélisation en effets fixes**

## **Résumé :**

---

L'immigration occupe depuis plusieurs années une place centrale dans les débats politiques et médiatiques en France, souvent associée à une hausse de la criminalité. Ce mémoire propose une analyse économétrique rigoureuse de cette relation à l'échelle départementale, sur la période 2006-2021, en mobilisant des données officielles du ministère de l'Intérieur et de l'INSEE. À l'aide de modèles en panel à effets fixes, intégrant des variables de contrôle et des erreurs clusterisées, l'étude évalue l'impact de la proportion d'immigrés sur différents types de criminalité. Les résultats montrent majoritairement une absence de lien statistiquement significatif entre immigration et criminalité. Toutefois, des effets variés apparaissent, à savoir, une baisse de certaines infractions complexes nécessitant une connaissance du système local (fraudes, plaintes, atteintes aux institutions), une hausse d'infractions mineures non violentes (vols sans violence), et une hausse des infractions diverses. Ces constats soulignent la nécessité de renforcer les politiques d'intégration, de promouvoir une meilleure sensibilisation aux études scientifiques rigoureuses, et de soutenir les initiatives visant à améliorer les conditions socio-économiques des populations immigrées.

---

**Mots clés :** Immigration, Criminalité, Données de Panel, Variables de contrôles, Erreurs standards clusterisées, RStudio

# **Analysis of the Effect of Immigration on Crime and Delinquency in French Departments between 2006 and 2021, Using a Fixed Effects Model**

## **Abstract:**

---

Immigration has occupied a central place in political and media debates in France for several years, often being associated with a rise in crime. This thesis offers a rigorous econometric analysis of this relationship at the departmental level over the period 2006-2021, using official data from the Ministry of the Interior and INSEE. Using fixed-effects panel models that include control variables and clustered standard errors, the study evaluates the impact of the proportion of immigrants on various types of crime. The results largely show no statistically significant link between immigration and crime. However, heterogeneous effects emerge: a decrease in certain complex offenses requiring knowledge of the local system (such as fraud, complaints, and offenses against institutions), an increase in minor non-violent offenses (such as non-violent theft), and a rise in miscellaneous offenses. These findings highlight the need to strengthen integration policies, promote greater awareness of rigorous scientific research, and support initiatives aimed at improving the socio-economic conditions of immigrant populations.

---

**Keywords:** Immigration, Crime, Panel Data, Control Variables, Clustered Standard Errors, RStudio

# Sommaire

<b>Glossaire sigles .....</b>	<b>2</b>
<b>Introduction .....</b>	<b>3</b>
<b>I - Analyse et revue de la littérature .....</b>	<b>6</b>
<b>II - Présentation des variables d'intérêt.....</b>	<b>9</b>
<b>II - Présentation des variables de contrôles.....</b>	<b>16</b>
<b>III - Modèles économétriques utilisées.....</b>	<b>19</b>
<b>IV - Méthodologie appliquée .....</b>	<b>24</b>
<b>V - Analyse des résultats.....</b>	<b>31</b>
<b>VI - Analyse des estimations des bases alternatives .....</b>	<b>37</b>
<b>VII - Résumés des résultats finaux.....</b>	<b>45</b>
<b>VIII - Discussion.....</b>	<b>46</b>
<b>Conclusion .....</b>	<b>50</b>
<b>Bibliographie, Sitographie, liens des bases de données.....</b>	<b>52</b>
<b>ANNEXES.....</b>	<b>56</b>
<b>Table des matières .....</b>	<b>72</b>

## Glossaire sigles

- INSEE : Institut nationale de la statistique et des études économiques
- INED : Institut national d'études démographiques
- IRIS : Ilots Regroupés pour l'Information Statistique
- SSMSI : Service statistique ministériel de la sécurité intérieure
- CSP : Catégories Socioprofessionnelles
- BIT : Bureau international du travail
- MCO : Moindres carrés ordinaires
- AIC : Akaike Information Criterion
- VIF : Variance Inflation Factor
- *SAPHIR* : Système d'analyse de la population par l'historique des recensements
- GMM : Generalized Method of Moments (Méthode des Moments Généralisée)
- DROM : Départements et Régions d'Outre-Mer

# Introduction

Depuis plusieurs décennies, l'immigration constitue un enjeu central des débats politiques et médiatiques en France. L'arrivée de populations étrangères est fréquemment perçue, comme une source potentielle de tensions sociales, économiques et sécuritaires. Parmi les idées reçues les plus persistantes figure l'association entre immigration et criminalité. Cette présumée causalité alimente régulièrement les discours populistes, parfois empreints de stigmatisation et influence certaines orientations politiques et législatives. Patrick Simon directeur de recherche à l'unité Migrations internationales et minorités à l'Ined, souligne que l'immigration est souvent évoquée comme un thème des transformations de la société, focalisant les préoccupations sur le logement, l'emploi, la laïcité et les inégalités.<sup>1</sup>

Lors des élections européennes de 2024, le Rassemblement national a obtenu 31,4 % des voix, en majorité grâce à une campagne centrée sur la question migratoire. Selon un sondage Ipsos, 43 % des électeurs français ont indiqué que l'immigration était un facteur décisif de leur vote, plaçant ce sujet derrière le pouvoir d'achat (45 %) mais devant la protection de l'environnement (27 %).<sup>2</sup>

Par ailleurs, le traitement médiatique des faits divers, souvent amplifié par certaines figures politiques, contribue à entretenir et renforcer les stéréotypes liés à l'immigration. Les données statistiques peuvent parfois être interprétées de manière biaisée, voire manipulées à des fins idéologiques ou électorales. Cela souligne la nécessité de mener une étude rigoureuse et objective sur les liens potentiels entre immigration et criminalité.<sup>3</sup>

En 2023, la France comptait 7,3 millions d'immigrés, soit 10,7 % de la population totale, un chiffre en constante augmentation depuis 1946.<sup>4</sup> Cette part importante de la population souligne l'intérêt de mieux comprendre les effets de l'immigration sur la société. La même année, les données du ministère de l'Intérieur ont révélé une hausse de 12 % des tentatives d'homicide et de 8 % des violences sexuelles. Bien que, ces augmentations doivent être interprétées avec prudence, car elles s'expliquent en partie

---

<sup>1</sup> Immigration : Le débat « peut être apaisé et dépolitisé par la connaissance et le dialogue »

<sup>2</sup> Elections européennes : comment la question migratoire a pesé sur le vote des Français

<sup>3</sup> Immigration et délinquance : attention aux interprétations trompeuses des chiffres du ministère de l'Intérieur – DE FACTO – Des clés pour mieux s'informer

<sup>4</sup> L'essentiel sur... les immigrés et les étrangers | Insee

par la libération de la parole des victimes et par l'amélioration de l'accueil dans les services de police, ce qui ne reflète pas nécessairement une augmentation objective de la violence.<sup>5</sup>

Dans ce contexte, la relation entre immigration et criminalité ne peut être réduite à une simple corrélation apparente. De nombreuses études empiriques soulignent la complexité de la relation, certaines concluant en l'absence de relation significative, tandis que d'autres mettent en évidence des effets spécifiques selon les types de criminalité, les profils sociodémographiques des immigrés ou les caractéristiques locales. Il est donc indispensable d'aborder cette question à l'aide d'une analyse rigoureuse, fondée sur des données objectives et des méthodes méticuleuses.

L'objectif de ce mémoire est d'analyser empiriquement si l'immigration constitue un facteur d'augmentation de la criminalité en France. Plus précisément, il s'agira d'évaluer l'effet de la proportion d'immigrés dans la population sur différents indicateurs de criminalité à l'échelle départementale, en s'appuyant sur des données couvrant la période de 2006 à 2021, en France métropolitaine.

Pour répondre à cette problématique l'étude vise à identifier l'existence d'un lien de causalité, et non d'une simple corrélation, entre immigration et criminalité. Une approche économétrique sur données de panel sera mobilisée, intégrant notamment des effets fixes, des variables de contrôles et des clusters « groupe », afin de contrôler l'hétérogénéité non observée entre départements et par année, d'éviter les biais de variables omises et de corriger la présence d'hétéroscédasticité et d'autocorrélation intra-groupe. Les données proviennent de sources officielles telles que le ministère de l'Intérieur et *data.gouv.fr*<sup>6</sup> pour les statistiques sur la criminalité et l'INSEE pour les données relatives à l'immigration ainsi que pour les variables de contrôle socio-économiques comme le sexe, l'âge, la catégorie socio-professionnelle et le niveau de vie.<sup>78</sup> L'utilisation de sources fiables et vérifiées est d'une importance capitale pour garantir la robustesse et la crédibilité des résultats, puisque dans le cadre d'une étude économétrique, la qualité des sources conditionne la validité des estimations et la pertinence des recommandations qui en

---

<sup>5</sup> Insécurité et délinquance en 2023 : bilan statistique et atlas départemental de la délinquance enregistrée | Ministère de l'Intérieur

<sup>6</sup> <https://www.data.gouv.fr/fr/>

<sup>7</sup> <https://www.insee.fr/fr/accueil>

<sup>8</sup> <https://www.interieur.gouv.fr/>



découlent. Il est donc essentiel de s'appuyer sur des bases de données officielles, des publications scientifiques reconnues, ainsi que sur des travaux d'experts afin d'assurer une analyse fiable et objectivement fondée.

Pour ce mémoire, nous débuterons par une revue de la littérature portant sur les relations entre immigration et criminalité. Nous présenterons ensuite le traitement des variables d'intérêt et l'implication des variables de contrôles socio-économiques. Puis, nous expliquerons le choix et le fonctionnement des modèles économétriques retenues. La méthodologie appliquée sera ensuite exposée, suivie des statistiques descriptives et de l'analyse de la validité des modèles estimés. Enfin, nous exposerons les résultats obtenus avant d'en discuter les principaux apports, les limites éventuelles et les perspectives qu'ils ouvrent. Il convient de noter qu'une section détaillant la construction des bases de données utilisées est disponible en annexe page 72.

## I - Analyse et revue de la littérature

La relation entre immigration et criminalité a fait l'objet de nombreuses études et d'une importante littérature empirique, principalement aux États-Unis, où les résultats convergents largement vers l'absence d'une corrélation positive entre ces deux phénomènes.

Les travaux pionniers de Butcher et Piehl (1998) utilisant une méthode en coupe transversale sur des données américaines pendant les années 1980, ont montré que les villes avec des taux de criminalité élevés ont tendance à avoir un grand nombre d'immigrants, mais que une fois contrôlé par les caractéristiques démographiques, les immigrants récents n'ont pas d'effet significatif sur les taux de criminalité. De plus, ils ont également montré que les jeunes nés à l'étranger sont statistiquement moins susceptibles d'être impliqués dans des activités criminelles que les jeunes natifs.<sup>9</sup> Cette approche a été approfondie par Rumbaut et Ewing en 2007 qui, analysant les données du recensement américain de 1980 à 2000, ont confirmé une corrélation négative entre immigration et criminalité au niveau des États-Unis. Ils ont notamment démontré que les immigrants, y compris ceux qui sont peu éduqués ou en situation irrégulière, ont des taux d'incarcération plus bas que les natifs.<sup>10</sup> Dans cette même perspective, Odabasi (2021) propose une analyse économétrique sur données de panel au niveau des comtés américains pour la période 2012–2015. En utilisant un modèle à effets fixes avec variables instrumentales, l'auteur distingue les impacts des migrations internes et internationales sur la criminalité. Les résultats indiquent que la migration interne est associée à une hausse des crimes violents et des délits contre les biens, tandis que la migration internationale n'a qu'un effet marginal et non systématique sur ces types de criminalité. Par ailleurs, l'étude met en évidence le rôle déterminant du chômage dans la variation des taux de criminalité.<sup>11</sup>

En Europe, plusieurs travaux aboutissent à des résultats similaires. Bell, Machin et Fasani (2013) ont adopté une approche en différences-en-différences (DiD) avec des effets fixes

---

<sup>9</sup> Butcher, K. F., & Piehl, A. M. (1998). Cross-city evidence on the relationship between immigration and crime. *Journal of Policy Analysis and Management*, 17(3), 457-493.

<sup>10</sup> Rumbaut, R. G., & Ewing, W. A. (2007). *The myth of immigrant criminality and the paradox of assimilation*. Immigration Policy Center.

<sup>11</sup> Odabasi, S. (2021). The economics of crime and immigration: A panel data analysis. *Pamukkale University Journal of Social Sciences Institute*.

ainsi qu'une approche par variables instrumentales, exploitant la répartition géographique quasi-aléatoire des demandeurs d'asile en Angleterre, puisque déterminée par des critères administratifs. Leur analyse ne met en évidence aucun effet statistiquement significatif de l'immigration sur les crimes violents ou les crimes de propriété, mais révèle une légère augmentation des délits mineurs dans les zones ayant accueilli davantage de demandeurs d'asile.<sup>12</sup> Plus récemment, Gehrsitz et Ungerer (2017) ont analysé l'afflux massif de demandeurs d'asile en Allemagne répartis entre les districts allemands selon des critères administratifs d'urgence tels que la capacité d'accueil. Exploitant une quasi-expérience naturelle en utilisant une spécification en panel avec effets fixes, ils ont trouvé que l'afflux de réfugiés en Allemagne n'a pas eu d'effets significatifs sur l'emploi des travailleurs natifs ni sur la criminalité. Par ailleurs, ils ont montré que la proximité directe aux réfugiés peut atténuer les attitudes négatives à leur égard.<sup>13</sup> De surcroît, Bianchi, Buonanno et Pinotti (2012) examinent l'impact causal de l'immigration sur la criminalité en Italie en exploitant les variations de la répartition des immigrés due aux quotas administratifs. À l'aide de modèles à effets fixes et d'une approche par variables instrumentales, ils montrent que l'immigration n'augmente pas significativement la criminalité violente, mais a un effet limité et positif sur certains délits mineurs comme les vols.<sup>14</sup> Dans une étude réalisée sur des données provinciales de 1999 à 2009 en Espagne et utilisant des modèles de panel sophistiqué comme le Système GMM qui permet de traiter les problèmes d'endogénéité, Alonso-Borrego, Garoupa et Vázquez ont montré que la forte vague d'immigration durant cette période, n'a pas causé d'augmentation significative de la criminalité pour les crimes violents mais qu'elle est liée à la hausse des délits mineurs et des crimes contre les biens. Ils soulignent que les caractéristiques spécifiques des immigrants, telles que l'origine, le niveau d'éducation, la langue et le genre, jouent un rôle déterminant dans ces relations.<sup>15</sup>

En France, cette thématique demeure moins exhaustive, en partie due aux limitations dans la collecte de données statistiques sur l'origine ethnique des individus. Jobard, Lévy,

---

<sup>12</sup> Bell, B., Fasani, F., & Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, 95(4), 1278-1290.

<sup>13</sup> Gehrsitz, M., & Ungerer, M. (2017). Jobs, crime, and votes: A short-run evaluation of the refugee crisis in Germany. *ZEW Discussion Paper*, 17-016.

<sup>14</sup> Bianchi, M., Buonanno, P., & Pinotti, P. (2012). *Do immigrants cause crime?* Journal of the European Economic Association, 10(6), 1318-1347.

<sup>15</sup> Alonso-Borrego, C., Garoupa, N., & Vázquez, P. (2012). Does immigration cause crime? Evidence from Spain. *American Law and Economics Review*, 14(1), 165-191.

Lamberth et Névanen (2012) ont analysé les pratiques de contrôle d'identité à Paris en 2007-2008, et ils ont observé que les individus perçus comme "noirs" ou "arabes" étaient significativement plus susceptibles d'être contrôlés que ceux perçus comme "blancs", même après avoir pris en compte des variables telles que l'âge, le sexe et le style vestimentaire. Relevant que la surreprésentation apparente des étrangers dans les données policières résultait largement de biais dans les contrôles d'identité et les pratiques policières.<sup>16</sup> De plus, l'étude de Aoki, Y., & Todo, Y. (2009) qui examine l'impact de l'immigration sur la criminalité en France par départements et utilisant des méthodes par effets fixes et variables de contrôles, révèlent que l'immigration seule n'augmente pas la criminalité en France. Cependant, les conditions économiques défavorables comme le chômage, peuvent accroître la propension des immigrants à commettre des crimes.<sup>17</sup>

Les différentes études et travaux empiriques concluent majoritairement à l'absence de lien systématique entre l'immigration et une hausse de la criminalité. Toutefois, certaines études relèvent un effet positif limité, notamment pour certains types de délits mineurs tels que les vols ou les infractions liées aux biens. Dans le contexte français, peu d'étude rigoureuse ont été menée, en grande partie en raison de contraintes liées à la disponibilité, la limitation et à la précision des données statistiques. Cette étude constitue ainsi une contribution méthodologique et empirique notable, en mobilisant des données de panel départementales sur plusieurs années. Les modèles à effet fixes et variables de contrôles utilisés permettent de contrôler les hétérogénéités inobservables constantes dans le temps, comme la culture locale ou l'efficacité des institutions, et permettent de mieux identifier des relations causales que dans les analyses purement transversales.

---

<sup>16</sup> Jobard, F., Lévy, R., Lamberth, J., & Névanen, S. (2012). Mesurer les discriminations selon l'apparence : une analyse des contrôles d'identité à Paris. *Revue française de sociologie*, 53(3), 423-451.

<sup>17</sup> Aoki, Y., & Todo, Y. (2009). Are immigrants more likely to commit crimes? Evidence from France. *Applied Economics Letters*, 16(15), 1537-1541.

## II - Présentation des variables d'intérêt

### A) Taux d'immigration

Selon l'INSEE, « la définition adoptée par le Haut Conseil à l'Intégration, un immigré est une personne née étrangère à l'étranger et résidant en France. Les personnes nées Françaises à l'étranger et vivant en France ne sont donc pas comptabilisées. Certains immigrés ont pu devenir Français, les autres restant étrangers. Les populations étrangère et immigrée ne se recoupent que partiellement : un immigré n'est pas nécessairement étranger et réciproquement, certains étrangers sont nés en France (essentiellement des mineurs). La qualité d'immigré est permanente : un individu continue à appartenir à la population immigrée même s'il devient Français par acquisition. C'est le pays de naissance, et non la nationalité à la naissance, qui définit l'origine géographique d'un immigré. »<sup>18</sup>

Le taux d'immigration est défini comme le rapport entre le nombre d'immigrés et la population totale d'un territoire et d'une année donnés. Cette mesure permet de standardiser l'information et de rendre les comparaisons pertinentes entre des départements de tailles démographiques différentes. La construction d'un taux plutôt qu'un nombre absolu facilite l'étude de l'impact relatif de la population immigrée sur les phénomènes étudiés.

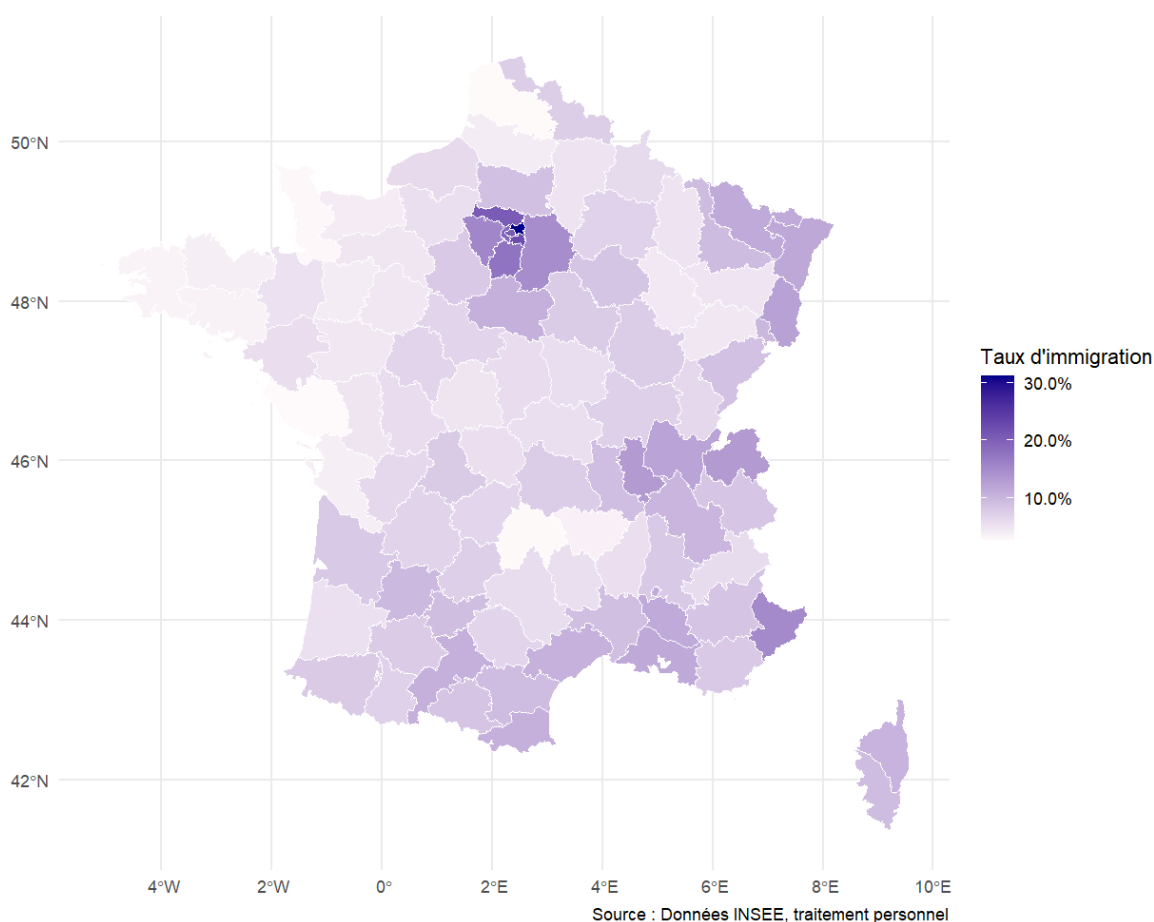
À titre d'illustration, en 2021 la population de la Loire-Atlantique (44) était estimée à 1 457 806 habitants, dont environ 81 558 immigrés, soit un taux d'immigration de 5,59 %. En 2006, ce taux s'élevait à 3,18 %, témoignant d'une hausse progressive de la part des immigrés dans la population du département et reflétant les dynamiques migratoires sur cette période. Cette évolution s'inscrit dans une tendance plus générale observée à l'échelle nationale (voir annexe 1). Ces tendances nationales et locales peuvent être visualisées à travers la carte ci-dessous (Figure 1), qui présente la répartition du taux d'immigration par département pour l'année 2021.<sup>19</sup>

---

<sup>18</sup> Définition - Immigré | Insee

<sup>19</sup> L'essentiel sur... les immigrés et les étrangers | Insee

Figure 1 : Répartition du taux d'immigration par département en 2021



Les données sur la population d'immigrés sont issues des estimations du recensement de la population réalisées par l'INSEE. Contrairement à un recensement exhaustif mené à une date unique sur tout le territoire, le dispositif actuel repose sur une collecte annuelle par sondage, mise en place depuis 2004. Les communes de moins de 10 000 habitants font l'objet d'un recensement exhaustif une fois tous les cinq ans, tandis que dans les communes de 10 000 habitants ou plus, un échantillon de 8 % des logements est enquêté chaque année. En agrégeant les résultats de cinq années consécutives, l'INSEE estime les caractéristiques de la population au 1er janvier de la 3<sup>ème</sup> année.<sup>20</sup> Ce mécanisme permet d'assurer la continuité des données dans le temps et une fiabilité élevée, bien qu'il repose sur des estimations. Les données diffusées ne sont donc pas parfaitement exactes, bien qu'elles respectent les standards de qualité définis par le code de bonnes pratiques de la statistique européenne.<sup>21</sup>

<sup>20</sup> Documentation complémentaire sur le recensement | Insee

<sup>21</sup> Conseils pour l'utilisation des résultats statistiques | Insee

Les informations sont recueillies via des questionnaires adressés aux habitants, qui collectent notamment le lieu de naissance et la nationalité à la naissance. Ces deux variables permettent d'identifier les immigrés au sens de la définition adoptée par l'INSEE. Les réponses sont géolocalisées à l'échelle infra-communale grâce au découpage en IRIS, qui représente l'unité de base pour la diffusion de données locales. Les IRIS sont en général constitués de 1 800 à 5 000 habitants et sont classés selon leur usage (habitat, activité, divers). Ils sont conçus pour être stables et comparables dans le temps, sauf en cas d'ajustements documentés.<sup>22</sup>

Dans le cadre de cette étude, les données issues des IRIS ont été agrégées au niveau départemental pour chaque année de 2006 à 2021, afin de construire la variable taux d'immigration ainsi que les différentes variables socio-économiques. Ce traitement permet de garantir la cohérence temporelle et spatiale des données, tout en facilitant leur exploitation dans un modèle économétrique. Bien que fondées sur des estimations, les données mobilisées ici sont considérées comme fiables, représentatives et rigoureusement encadrées méthodologiquement.

Il est également important de souligner que les données de l'INSEE ne permettent ni de distinguer les immigrés selon leur statut administratif (régulier ou irrégulier), ni d'identifier les situations de séjour temporaire ou clandestine. La définition statistique retenue par l'INSEE englobe uniquement les individus ayant une résidence stable en France, quelle que soit leur nationalité actuelle, mais n'inclut pas les personnes nées en France de parents immigrés, ni les étrangers en situation irrégulière s'ils ne sont pas recensés.

Cette limitation méthodologique a des implications importantes pour l'analyse, notamment si l'on cherche à explorer l'impact de l'immigration sur la criminalité. D'une part, les représentations sociales de l'immigration incluent souvent des dimensions qui échappent à cette définition statistique, comme l'immigration irrégulière. D'autre part, si certains discours associent la criminalité à une immigration clandestine ou précaire, cette réalité n'est pas observable directement dans les données mobilisées ici. Il convient donc d'interpréter les résultats avec prudence, en gardant à l'esprit que la variable utilisée

---

<sup>22</sup> Définition - IRIS | Insee

mesure uniquement l'immigration "administrativement visible" et statistiquement recensée.

## **B) Taux de criminalité**

La criminalité désigne l'ensemble des actes punis par la loi pénale. Elle représente à la fois un enjeu central de sécurité publique et un indicateur révélateur des dynamiques sociales d'un territoire.<sup>23</sup> En France, les préoccupations liées à l'insécurité ont occupé une place croissante dans les débats publics au cours des dernières décennies, alimentées par l'évolution des faits enregistrés, mais aussi par la médiatisation des actes violents et des incivilités.

Depuis plusieurs années, le paysage criminel français a connu d'importantes transformations. Alors que les atteintes aux biens (cambriolages, vols, etc.) ont tendance à diminuer depuis le début des années 2000, les violences physiques, en particulier les violences intrafamiliales et les agressions, sont en augmentation constante. Le ministère de l'Intérieur note ainsi que les violences contre les personnes ont plus que doublé en vingt ans. Cette hausse s'explique en partie par une plus grande propension des victimes à porter plainte, mais aussi par une évolution des comportements violents (voir annexe 2).<sup>24</sup>

Dans cette étude, les taux de criminalité utilisés sont calculés à partir des données de crimes et délits enregistrés par les forces de police et de la gendarmerie nationales depuis janvier 1996, regroupées mensuellement par département dans la base administrative historique dite « État 4001 ». <sup>25</sup> Ces données sont issues du ministère de l'Intérieur, puis du SSMSI à partir de 2014. Le SSMSI supervisé par l'Autorité de la statistique publique, en tant que service statistique public applique des normes strictes en matière de neutralité,

---

<sup>23</sup> Définition - Criminalité | Insee

<sup>24</sup> Insécurité et délinquance en 2023 : bilan statistique et atlas départemental de la délinquance enregistrée | Ministère de l'Intérieur

<sup>25</sup> Chiffres départementaux mensuels relatifs aux crimes et délits enregistrés par les services de police et de gendarmerie depuis janvier 1996 - data.gouv.fr



d'indépendance, de qualité méthodologique et d'accessibilité concernant les données et les résultats utilisés.<sup>26</sup>

La base « État 4001 » recense les infractions pénales enregistrées pour la première fois par les services de sécurité, tels que les crimes, délits, atteintes aux biens ou aux personnes, etc. Classés selon 107 catégories appelées « index », ces faits peuvent avoir été portés à la connaissance des autorités par le biais de plaintes, de signalements, de constatations en flagrant délit, ou d'initiatives des forces de l'ordre.<sup>27</sup>

Le taux de criminalité est ainsi défini comme le rapport entre le nombre total d'infractions enregistrées (hors contraventions) et la population totale d'un département donné, pour une année donnée, et pour 100 000 habitants. Le taux pour 100 000 habitants constitue une mesure standard largement utilisée dans la littérature. Il permet de normaliser les données et de comparer l'intensité relative de la criminalité entre des territoires de tailles démographiques différentes. À noter que certaines infractions, notamment celles ayant des unités de compte différentes (victime, victime entendue, auteur, véhicule, etc.), ne peuvent pas toujours être agrégées sans perte de rigueur. Le traitement statistique adopté ici veille à ne considérer que les catégories homogènes et comparables dans le temps.

Pour ce faire, les index ont été regroupés selon leur unité de compte afin de construire huit indicateurs distincts représentant différents types de criminalité (voir Tableau 1). Ce regroupement permet de garantir la cohérence des comparaisons dans le temps et entre départements, en évitant les biais liés à l'agrégation d'infractions de nature hétérogène.

Tableau 1 : Présentation des variables relatives aux taux de criminalités

Catégories	Description	Exemple d'index inclus
<i>taux crime chèques</i>	Falsifications et usages frauduleux de chèques	Falsifications et usages de chèques volés
<i>taux crime plaintes</i>	Infractions signalées par des plaignants	Escroqueries et abus de confiance,

<sup>26</sup> Insécurité et délinquance en 2023 : bilan statistique et atlas départemental de la délinquance enregistrée | Ministère de l'Intérieur

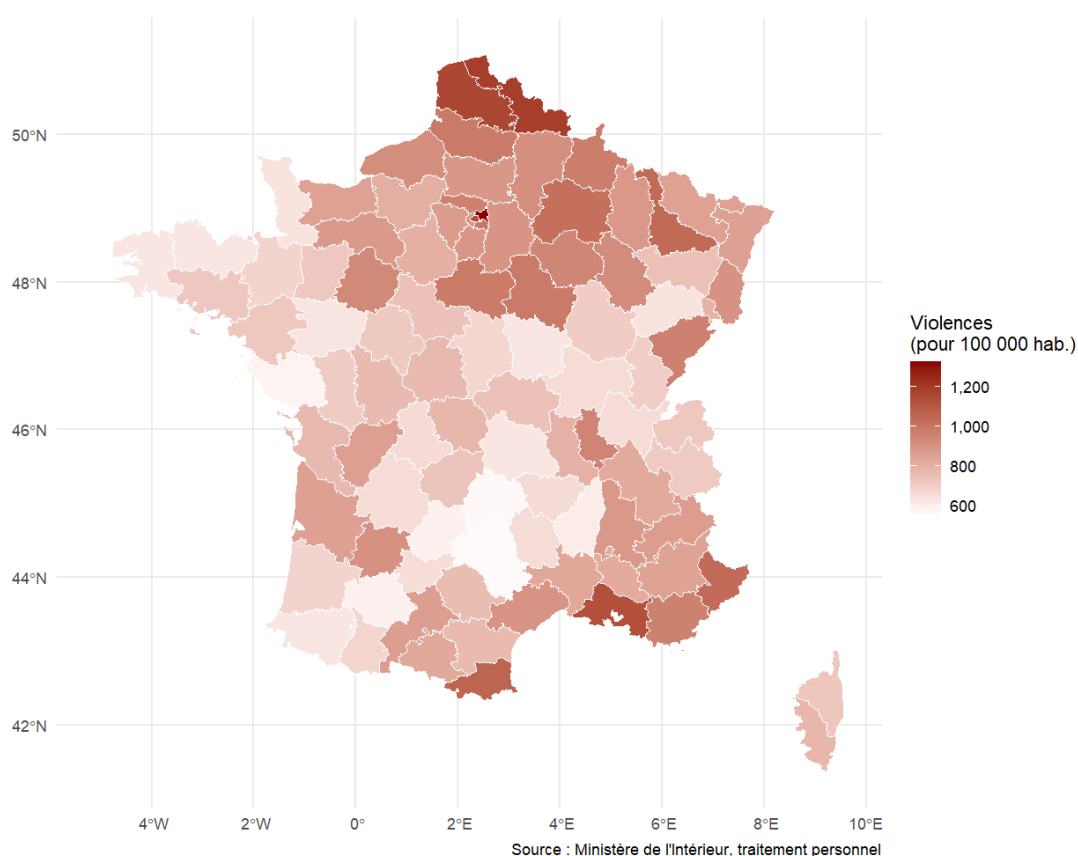
<sup>27</sup> Bases statistiques communale, départementale et régionale de la délinquance enregistrée par la police et la gendarmerie nationales (base) - data.gouv.fr

		Autres destructions et dégradations de biens publics, etc.
<i>taux crime vols vehicules</i>	Vols ou dégradations sur des véhicules	Vols d'automobiles, vols à la roulotte, vols d'accessoires sur véhicules à moteur immatriculés, etc.
<i>taux crime stupefiants</i>	Infractions liées aux stupéfiants et enregistrée comme « auteur »	Usage de stupéfiants, Usage-revente de stupéfiants, Non-versement de pension alimentaire, etc.
<i>taux crime vols victime</i>	Vols simples, à l'étalage ou à la tire	Vols à la tire, Vols à l'étalage, Autres vols simples, etc.
<i>taux crime violence</i>	Violences physiques et psychologiques graves	Coups et blessures volontaires suivis de mort, Viols sur des majeur(e)s, Homicides pour d'autres motifs, Harcèlements sexuels, etc.
<i>taux crime autres</i>	Infractions diverses enregistrées comme "procédures"	Proxénétisme, Séquestrations, Travail clandestin, Banqueroutes, Faux documents d'identité, Attentats à l'explosif, etc.
<i>taux crime infractions pub</i>	Infractions contre des institutions ou des établissements publics	Cambriolages de locaux industriels, Prises d'otages à l'occasion de vols, etc.

Bien que ces données soient jugées fiables, il convient de souligner que ces statistiques ne reflètent pas l'ensemble de la délinquance et de la criminalité réelle. Elles ne capturent pas les délits non déclarés, ne permettent pas clairement de distinguer la gravité des faits, elles sont sensibles à la proportion de victimes qui portent plainte et à l'intensité de l'activité policière. Elles ne mesurent donc que la criminalité "enregistrée", qui peut être influencée par des facteurs institutionnels, sociétaux ou médiatiques. Pour une vision plus complète, l'INSEE et le SSMSI recommandent généralement de compléter ces données par des enquêtes de victimisation. Ces dernières ne sont pas mobilisées dans le cadre de ce travail en raison de leur non-disponibilité à l'échelle départementale.<sup>28</sup>

Enfin, la carte ci-dessous illustre la répartition de la criminalité en France métropolitaine, elle présente les taux de crimes violents enregistrés par départements en 2021, exprimé pour 100 000 habitants. Cette visualisation met en évidence les disparités géographiques en matière de violences physiques et psychologiques, et constitue un premier aperçu des zones où la criminalité violente est la plus concentrée (Figure 2).

Figure 2 : Taux de crimes violents par département en 2021 (pour 100 000 habitants)



<sup>28</sup> Indicateurs annuels de la victimation et du sentiment d'insécurité issus des enquêtes Cadre de Vie et Sécurité (CVS) - data.gouv.fr

## II - Présentation des variables de contrôles

Afin d'isoler au mieux l'effet de l'immigration sur les taux de criminalité, il est essentiel d'introduire des variables de contrôle dans les différentes régressions. Ces variables permettent de tenir compte de facteurs structurels ou conjoncturels susceptibles d'influencer simultanément la criminalité et le taux d'immigration. En d'autres termes, elles permettent de réduire le biais d'omission de variables pertinentes, et donc d'obtenir une estimation plus robuste et plus crédible de la relation étudiée.<sup>29</sup> Par exemple, des départements à forte immigration pourraient aussi présenter une structure socio-économique ou démographique particulière (jeunes adultes, chômage élevé, etc.) qui contribue elle-même à expliquer les niveaux de criminalité. Ne pas en tenir compte exposerait l'analyse à des biais d'endogénéité ou de confusion.<sup>30,31</sup>

Ainsi, cette étude mobilise trois grandes catégories de variables de contrôle : deux variables démographiques, une variable socio-professionnelle, et trois variables économiques.

### a) Variables démographiques

La part des hommes : De nombreuses études ont montré que les hommes, en particulier les jeunes hommes, sont statistiquement plus impliqués dans les actes de délinquance.<sup>32</sup> La variable utilisée ici mesure la proportion d'hommes dans la population totale d'un département pour une année donnée.\*

Tranche d'âge : La criminalité varie fortement selon l'âge, les 15-29 ans concentrent la majorité des actes de délits et de criminalité, tandis que les jeunes enfants ou les personnes âgées sont rarement auteurs ou victimes de délits. En intégrant les parts de

---

<sup>29</sup> Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach*. South-Western Cengage Learning.

<sup>30</sup> Endogénéité - SLPV analytics

<sup>31</sup> Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica*, 51(2), 277-304.

<sup>32</sup> Pons, S. (2017). Les crimes ont-ils un genre ? Étude statistique comparée de la criminalité masculine et féminine en Haute-Garonne au XIXe siècle. *Les Cahiers de Framespa*, 25.

\* La part des femmes n'a pas été incluse séparément dans les régressions car elle est parfaitement corrélée à la part des hommes ce qui crée un problème de colinéarité parfaite. En économétrie, cela conduit à la suppression automatique de cette variable par les logiciels de régression, car l'information apportée par l'une est strictement redondante avec l'autre.

population par tranches d'âge tous les 15 ans, on tient compte de cette structure démographique différenciée.<sup>3334</sup>

## **b) Variable socio-professionnelle**

Catégories socioprofessionnelles (CSP) : Le profil professionnel des habitants d'un territoire reflète de nombreuses inégalités. Que ce soient des inégalités économiques, d'insertion, de précarité ou de niveau d'éducation, ces facteurs peuvent exercer une influence sur les comportements criminels.<sup>35</sup> Par exemple, les ouvriers ou les employés sont plus exposés à certaines formes de criminalité que les cadres.<sup>36</sup> Cette variable est représentée en huit catégories socio-professionnelles, également en part de la population totale.

## **c) Variables économiques**

Taux de chômage : Le lien entre le chômage et la criminalité est complexe, mais largement documenté.<sup>37</sup> Certaines recherches suggèrent qu'un chômage élevé, en particulier chez les jeunes, peut accroître les incitations à commettre des délits, faute d'opportunités économiques légales.<sup>38</sup> Ici, le chômage est défini selon la méthodologie du BIT, appliquée en France par l'INSEE. Cette définition repose sur des critères précis, indépendants des régimes d'indemnisation ou des dispositifs administratifs, garantissant une mesure stable, harmonisée dans le temps et entre pays.<sup>39</sup> Le taux de chômage correspond donc au rapport entre le nombre de chômeurs et la population active totale (employés + chômeurs), ce qui est la mesure la plus courante pour analyser l'insertion sur le marché du travail (voir annexe 3).

---

<sup>33</sup> Farrington, D. P. (1986). Age and crime. *Crime and Justice*, 7, 189–250.

<sup>34</sup> Grogger, J. (1998). *Market wages and youth crime*. *Journal of Labor Economics*, 16(4), 756–791.

<sup>35</sup> Lagrange, Hugues. "Crime et conjoncture socio-économique." *Revue française de sociologie* 42, no. 1 (2001): 57–79.

<sup>36</sup> Fougère, D., Kramarz, F., & Pouget, J. (2005). *L'analyse économétrique de la délinquance : Une synthèse de résultats récents*. *Revue française d'économie*, 19(3), 3–42.

<sup>37</sup> Fougère, D., Pouget, J., & Kramarz, F. (2009). Youth Unemployment and Crime in France. *Journal of the European Economic Association*, 7(5), 909–938.

<sup>38</sup> Raphael, S., & Winter-Ebmer, R. (2001). *Identifying the effect of unemployment on crime*. *Journal of Law and Economics*, 44(1), 259–283.

<sup>39</sup> L'essentiel sur... le chômage | Insee

Niveau de vie médian : La médiane du niveau de vie, permet d'évaluer le niveau économique moyen d'un département. Un niveau de vie plus élevé peut être associé à des ressources institutionnelles plus importantes, à une meilleure prévention sociale, ou à des environnements moins criminogènes.<sup>40</sup>

Rapport interdécile (D9/D1) : Il s'agit d'un indicateur d'inégalités de revenus au sein de la population. Il est calculé en divisant le niveau de vie au 9e décile (D9, soit le seuil au-dessus duquel se situent les 10 % les plus riches) par celui du 1er décile (D1, soit le seuil en dessous duquel se situent les 10 % les plus pauvres). En d'autres termes, il mesure combien les 10 % les plus aisés gagnent par rapport aux 10 % les plus modestes. Par exemple, un rapport D9/D1 de 3,5 signifie que les plus riches ont un revenu 3,5 fois supérieur à celui des plus pauvres.<sup>41</sup> De nombreuses recherches ont mis en évidence une corrélation positive entre les inégalités de revenu et certains types de criminalité, en particulier les crimes contre les biens (vols, cambriolages, etc.).<sup>4243</sup>

---

<sup>40</sup> Chalfin, A., & McCrary, J. (2017). *Criminal deterrence: A review of the literature*. Journal of Economic Literature, 55(1), 5–48.

<sup>41</sup> Indicateur d'inégalité : le rapport interdécile - La finance pour tous

<sup>42</sup> Freeman, R. (1999). The economics of crime. Handbook of Labor Economics, 3, 3529–3571.

<sup>43</sup> Fajnzylber, P., Lederman, D., & Loayza, N. (2002). *Inequality and violent crime*. Journal of Law and Economics, 45(1), 1–39.

### III – Modèles économétriques utilisées

L'analyse empirique de cette étude repose sur l'estimation de modèles en données de panel, intégrant des effets fixes ainsi que des variables de contrôle, avec des erreurs standards clusterisées au niveaux départemental et annuel afin de corriger l'hétéroscédasticité et la dépendance intra-groupe.<sup>4445</sup> L'objectif est de s'éloigner d'une simple corrélation descriptive pour s'approcher d'une relation causale robuste entre immigration et criminalité.

Le modèle à effets fixes est une méthode économétrique fréquemment utilisée dans les analyses en données de panel. Une analyse en données de panel consiste à analyser des données observées de plusieurs individus sur plusieurs années, dans notre cas cela correspond à analyser l'impact du taux d'immigration pour l'ensemble des départements de France métropolitaine sur 16 ans (de 2006 à 2021). De ce fait, une observation correspond à un département pour une années donnée.

Ce type de modèle permet de contrôler les hétérogénéités non observées qui sont constantes dans le temps, qui peuvent varier entre les individus ainsi que les chocs communs affectant tous les individus à une même date. L'hétérogénéité non observée par individus correspond à des caractéristiques propres à chaque unité d'observation (ici, les départements) qui influencent le phénomène étudié, mais qui ne sont pas mesurées ni incluses dans les données.

Dans cette étude les caractéristiques inobservées par département peuvent inclure la qualité des institutions locales, les politiques d'inclusions et d'intégration des immigrés, l'histoire migratoire et la géographie des départements, les comportements sociaux vis-à-vis de l'immigration, mais aussi le niveau de coopération entre forces de l'ordre et population, les caractéristiques liées à la criminalité et la fermeté judiciaire ou encore les dynamiques culturelles spécifiques, etc.

De même, l'hétérogénéité non observée par année, correspond à des événements ou chocs macroéconomiques, politiques ou sociaux qui affectent simultanément tous les

---

<sup>44</sup> Arellano, M. (1987). Practitioners' corner: Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, 49(4), 445-450.

<sup>45</sup> White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.

départements. Par exemple, la pandémie de COVID-19 a bouleversé le fonctionnement des institutions judiciaires, les mobilités humaines et les dynamiques sociales, ce qui a pu influencer à la fois les niveaux de criminalité et les flux migratoires. D'autres exemples qui incluent des réformes nationales en matière de sécurité ou d'immigration, des crises économiques majeures comme la crise des subprimes (2008)<sup>46</sup>, des changements de gouvernement, ou encore des événements exceptionnels comme des attentats ou des mouvements sociaux d'ampleur nationale, sont susceptibles de modifier temporairement les comportements et les politiques nationales.

L'ensemble de ces caractéristiques inobservées au niveaux nationales et locales constituent des biais potentiels pouvant influencer les taux d'immigration et les taux de criminalité, ce qui rend inexploitable les résultats des estimations. L'objectif est d'isoler au maximum tous les facteurs externes, afin de n'obtenir que l'effet réel de l'immigration sur la criminalité.

En pratique, les effets fixes par département et par année permettent de neutraliser ces hétérogénéités non observées, qui sont deux sources potentielles de biais dans l'estimation de l'effet des variables explicatives. Pour ce faire, on introduit des variables indicatrices nommé « dummies » (ou variables muettes), qui vont absorber ces hétérogénéités.\* Les coefficients des variables muettes par départements représentent donc les différences moyennes entre chaque département et un département de référence. Et, les coefficients des variables muettes par années représentent les différences moyennes entre chaque année par rapport à une année de référence.

Le modèle s'écrit :

$$Y_{dt} = \alpha + \beta X_{dt} + \gamma_d + \delta_t + \varepsilon_{dt}$$

où :

- $Y_{dt}$  est la variable dépendante (taux de criminalité) observée dans le département d à l'année t.

---

<sup>46</sup> Les migrations des jeunes Européens après la crise de 2008 - Séminaires et colloques de l'Ined - Rencontres scientifiques - Ined - Institut national d'études démographiques

\* Ces variables prennent la valeur 1 lorsque l'observation appartient à une catégorie (par exemple, un département donné ou une année donnée), et 0 sinon. Leur inclusion revient à estimer un coefficient de variation spécifique pour chaque département et chaque année. Ce qui permet de contrôler les différences inobservables constantes dans le temps entre départements et les chocs communs à toutes les unités survenant une même année.



- $X_{dt}$  est la variable explicative d'intérêt (taux d'immigration).
- $\alpha$  est le terme constant du modèle (ordonnée à l'origine).
- $\beta$  est le coefficient à estimer qui mesure l'effet moyen de  $X_{dt}$  sur  $Y_{dt}$ , une fois contrôlé par les effets fixes. C'est l'effet de l'immigration sur les taux de criminalité.
- $\gamma_d$  sont les effets fixes propres à chaque département, capturant toutes les caractéristiques constantes dans le temps (politiques locales, comportements sociaux, spécificités culturelles, etc.).
- $\delta_t$  sont les effets fixes propres à chaque année, capturant les chocs communs à tous les départements (réformes nationales, chocs économiques, pandémie, etc.).
- $\varepsilon_{dt}$  est l'erreur aléatoire restante qui n'est pas capturée par le modèle, supposée centrée et non corrélée avec les variables explicatives.

L'inclusion de  $\gamma_d$  et  $\delta_t$  permet donc d'isoler l'effet de  $X_{dt}$  en utilisant uniquement la variation intra-départementale dans le temps, tout en contrôlant pour les différences structurelles entre départements et les tendances globales au fil des années.

Ce modèle est considéré comme robuste pour établir une relation causale sous certaines conditions, notamment si l'on suppose que les variables explicatives sont exogènes une fois que l'on a inclut les effets fixes. Il réduit les risques d'omission de variables corrélées avec les variables explicatives constante dans le temps, ce qui améliore la validité interne de l'analyse.<sup>47</sup> Cependant, il ne contrôle pas les variables omises qui varient dans le temps et qui pourraient encore biaiser les résultats si elles sont corrélées aux variables explicatives. En ce sens, il ne prouve pas à lui seul une causalité mais constitue une étape rigoureuse vers une identification plus crédible qu'un simple modèle de régression en coupe transversale. Enfin, le modèle à effets fixes peut souffrir d'une perte d'efficacité statistique si la variation temporelle des variables explicatives est faible.

C'est pourquoi, l'ajout de variable contrôle est nécessaire dans l'analyse. Ces variables permettent de capturer les variations temporelles susceptibles d'impacter simultanément la variable dépendante et la variable explicative. Leur inclusion renforce la validité interne du modèle en limitant le biais d'endogénéité dû à des variables omises. Autrement dit, les variables de contrôles isolent l'effet propres du taux d'immigration en tenant compte d'autres facteurs pouvant expliquer les taux de criminalité.

---

<sup>47</sup> Chytilova, H. (2015). Internal and external validity in experimental economics.

Le modèle s'écrit alors :

$$Y_{dt} = \alpha + \beta X_{dt} + \theta Z_{dt} + \gamma_d + \delta_t + \varepsilon_{dt}$$

où :

- $Z_{dt}$  représente les variables de contrôles qui varient par département et par année (ex : taux de chômage, niveau d'éducation, genre, âge, revenu médian, etc.).
- $\theta$  est le coefficient d'estimation des variables de contrôles qui mesure l'effet moyen de  $Z_{dt}$  sur  $Y_{dt}$ , une fois contrôlé par les effets fixes.

Toutefois, l'inclusion des variables de contrôles doit être fait avec précaution. Une variable de contrôle qui serait corrélé avec le taux d'immigration mais aussi causé par celui-ci est une variable endogène, c'est-à-dire corrélé au terme d'erreur, ce qui peut introduire de nouveaux biais et nuire à l'interprétation causale.<sup>48</sup>

Pour cette étude le modèle final s'écrit donc comme suit :

$$TAUX\_CRIMINALITE_{dt} = \alpha + \beta * TAUX\_IMMIGRATION_{dt} + \theta_1 * PART\_HOMMES_{dt} + \theta_2 * TRANCHES\_AGES_{dt} + \theta_3 * CATEGORIES\_SOCIO\_PROFESSIONELLES_{dt} + \theta_4 * TAUX\_CHOMAGE_{dt} + \theta_5 * MEDIANE\_NIVEAU\_DE\_VIE_{dt} + \theta_6 * RAPPORT\_INTERDECILE_{dt} + \gamma_d + \delta_t + \varepsilon_{dt}$$

En sachant que les erreurs standards sont corrigées par le recours à des erreurs standards robustes clusterisées (groupées), à la fois au niveau des départements et des années. Cette double clusterisation permet de prendre en compte l'hétéroscédasticité et l'autocorrélation des erreurs à l'intérieur de chaque département au fil du temps. L'intérêt de cette approche est d'éviter la sous-estimation de la variance des erreurs qui pourrait résulter de ces dépendances intra-groupes. Ignorer la possible autocorrélation ou variances non constantes des erreurs entres groupes peut biaiser les erreurs standards à la baisse et conduire à des déductions statistiques trop optimistes (notamment à des p-values trop faible). On risque alors de conclure à tort, à un effet significatif de l'immigration sur la criminalité alors qu'il n'en existe pas. La prise en compte simultanée des clusters par département et par année améliore ainsi la robustesse des tests

---

<sup>48</sup> Lechner, M. (2007). A note on endogenous control variables in causal studies. *SI/AW*.

statistiques et renforce la crédibilité des résultats empiriques sans pour autant impacter les résultats des estimateurs. Cette méthode est largement adoptée par la littérature et permet de valider les hypothèses statistiques classiques d'homoscédasticité et d'indépendance des erreurs ((Wooldridge, 2010) et Angrist et Pischke, 2009)).<sup>4950</sup> C'est donc ce modèle final qui sera retenu pour l'analyse et l'interprétation des résultats.

---

<sup>49</sup> Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.

<sup>50</sup> Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

## **IV - Méthodologie appliquée**

Cette section décrit les méthodes d'estimation et les bases de données mobilisées afin d'assurer la robustesse et la fiabilité des résultats. Elle présente ensuite les statistiques descriptives relatives aux variables du modèle final. Enfin, elle expose les différents tests de validité appliqués, destinés à garantir la pertinence et la rigueur de l'interprétation des résultats économétriques.

### **a) Méthodes d'estimations et bases de données**

Avant toute chose, il est judicieux de préciser que l'ensemble des résultats présentés dans cette étude concerne exclusivement la France métropolitaine. Les départements d'outre-mer Guadeloupe (971), Martinique (972), Guyane (973), La Réunion (974) et Mayotte (976) n'ont pas été inclus dans l'analyse en raison de la faiblesse ou de l'indisponibilité des données statistiques les concernant.

Afin de garantir la cohérence et la robustesse des résultats, l'analyse a été soumise à plusieurs méthodes d'estimations différentes qui permettent d'identifier la variabilité des résultats et leur stabilité. La régression finale repose sur un modèle à effets fixes départementaux et annuels, estimé par la méthode des MCO, avec variables de contrôles et erreurs standards clusterisées.

Dans un premier temps, une régression simple a été estimée afin d'examiner la corrélation simple entre le taux d'immigration et les taux de criminalité. Cependant, cette approche naïve ne permet pas d'isoler l'effet propre de l'immigration, car elle omet des variables importantes et ne tient pas compte des caractéristiques spécifiques à chaque département. Elle est donc sujette à un biais d'omission et ne peut conduire à une interprétation causale.

En complément, d'autres régressions ont été menées, certaines n'intègrent que les variables de contrôle, d'autres mobilisent des effets fixes simples (uniquement par département ou uniquement par année), ou encore des effets fixes combinés avec erreurs clusterisées uniquement selon l'une ou l'autre dimension. Ces différentes spécifications permettent de tester la sensibilité des résultats à l'inclusion des effets fixes, des variables

de contrôles et au traitement des erreurs. L'objectif est de vérifier si l'effet estimé de la variable d'intérêt (le taux d'immigration) reste significatif et cohérent selon la spécification retenue. Cette stratégie méthodologique vise à s'assurer que les conclusions de l'étude ne dépendent pas de manière critique d'un choix arbitraire de modèle ou d'une spécification.

Par ailleurs, une procédure de sélection pas à pas (stepwise), fondée sur la minimisation du critère AIC, a été appliquée au modèle final afin d'optimiser le compromis entre qualité d'ajustement et complexité du modèle. Cette méthode permet de retenir les variables les plus pertinentes tout en évitant le sur-ajustement.<sup>51</sup>

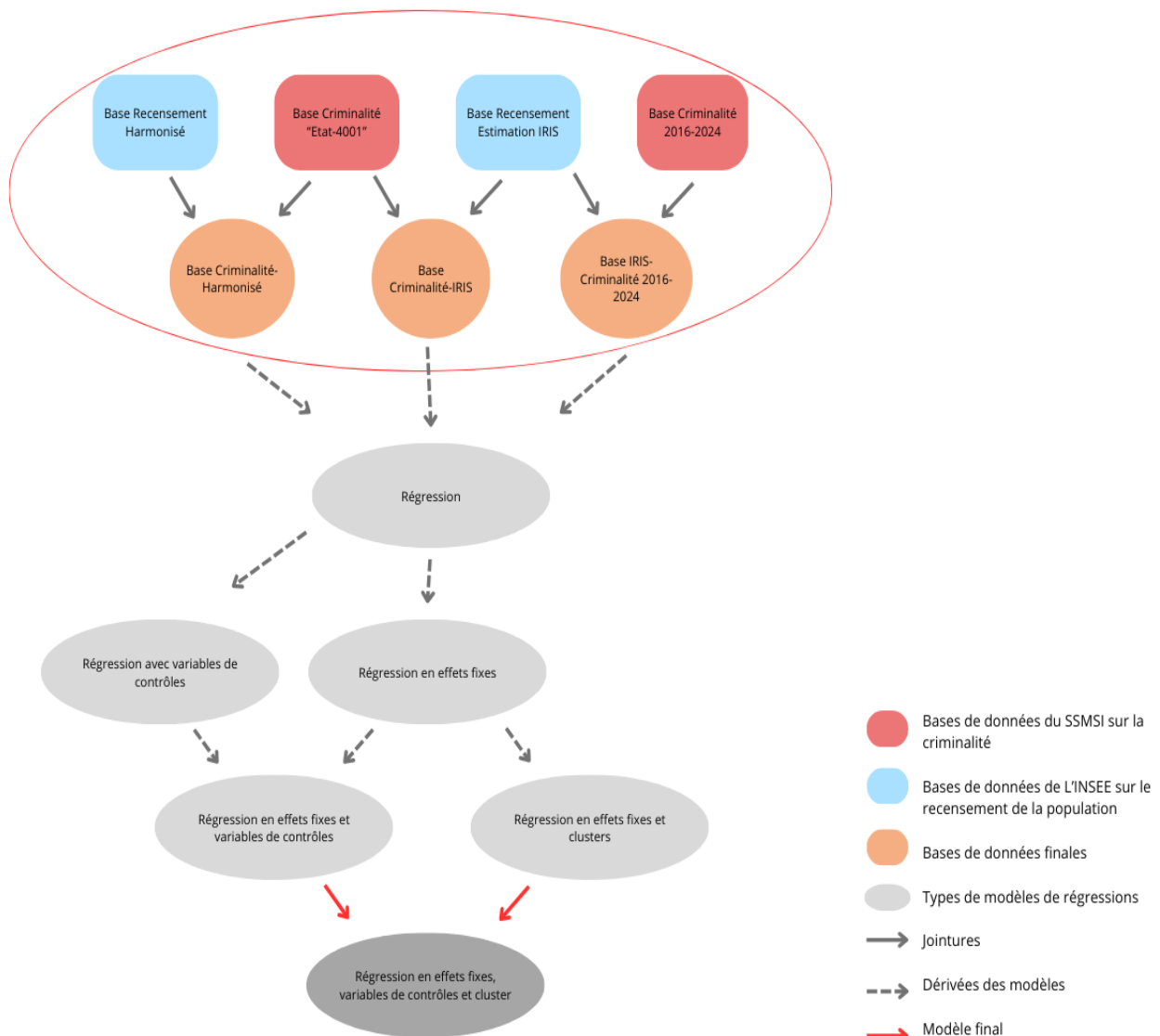
Deux autres bases de données alternatives, intégrant des informations de nature relativement différente, ont également été construites dans une démarche comparative visant à évaluer la robustesse des estimations et à confronter les résultats obtenus selon les sources et les définitions. La première base s'appuie sur les données du recensement harmonisé couvrant quatre années de données, et constitué avec les mêmes indicateurs de criminalité présenté ici. Bien que cette base de données soit riche en informations démographiques, sa couverture temporelle limitée ne permet pas d'identifier de manière certaine une relation causale. La seconde base mobilise des indicateurs de criminalité alternatifs, élaborés par le SSMSI et couvrant la période de 2016 à 2021. Celle-ci mobilise les mêmes données d'estimation du recensement présentés ici et est également intéressante au vu de la construction plus fine des catégories de criminalité, différenciant les types de victimes ou la localisation des faits. Pour en savoir davantage sur la construction et le traitement des bases de données veuillez-vous référer à l'annexe 25 page 72.

Les résultats issus de ces deux bases de données complémentaires seront présentés dans la section consacrée aux bases et techniques alternatives, et serviront à approfondir l'analyse principale. Enfin, la figure 3 ci-dessous présente une synthèse des différentes méthodologies employées ainsi que des bases de données élaborées au cours de l'analyse.

---

<sup>51</sup> Wooldridge, J. M. (2010). *Introductory Econometrics: A Modern Approach*. Cengage.

**Figure 3 : Organigramme des différentes bases de données et modèles utilisés**



Source : Canva, Auteur

Le modèle final à effets fixes départementaux et annuels, estimé par la méthode des MCO, avec variables de contrôles et erreurs standards doublement clusterisées, a été privilégié grâce à la richesse des informations qu'il fournit, sa capacité à neutraliser l'hétérogénéité non observée géographique et temporelle, ainsi que pour sa capacité globale d'ajustement mesurée par les indicateurs de performance tels que le  $R^2$  global, le  $Within R^2$  et le critère AIC. Ces derniers permettent d'évaluer respectivement la part de la variance totale et la

variance intra-départementale expliquées par le modèle, ainsi que de comparer la qualité de l'ajustement et la complexité de celui-ci.

## **b) Analyse descriptive des variables des modèles finaux**

Avant d'interpréter les résultats issus des modèles économétriques finaux, il est essentiel de comprendre et d'appréhender la structure des données mobilisées dans cette étude. Cette section est ainsi consacrée à une analyse descriptive des variables.

Pour commencer, la variable taux d'immigration exprimée en pourcentage, présente une moyenne de 7,389 % et une médiane à 6,436 % sur l'ensemble des départements et la période étudiée (voir annexe 7). Cette différence entre la moyenne et la médiane indique une asymétrie à droite de la distribution, due à la présence de départements affichant des taux particulièrement élevés à certaines périodes. L'annexe 4, qui présente l'histogramme et la boîte à moustaches de cette variable, met en évidence une distribution relativement normale, bien que contenant plusieurs valeurs extrêmes. À titre d'exemple, le département de la Seine-Saint-Denis (93) enregistre des taux d'immigrations qui évoluent de 26,47 % en 2006 et atteignant 31,14 % en 2021. Bien que ces valeurs soient élevées, elles ne peuvent être considérées comme atypiques ni aberrantes car elles traduisent une réalité locale qui mérite d'être intégrée à l'analyse.

Ensuite, les différents indicateurs de criminalité suivent une loi normale et présentent également des valeurs extrêmes, pour chaque indicateur ces valeurs représentent des départements avec plus de criminalité sur une période (voir annexe 5 et 6). Par exemple, pour l'indicateur de criminalité « *taux\_crime\_vols\_victimes* » qui correspond au taux de vols simple, à l'étalage ou à la tire, calculé à partir du nombre de victimes ayant déclarées ces infractions, met en évidence de fortes disparités territoriales. En 2019, le département de Paris (75) enregistre un maximum de 6103,6 vols pour 100 000 habitants, tandis que le minimum observé est de 317,8 vols pour 100 000 habitants dans le département de la Creuse (23) en 2010. La moyenne nationale sur l'ensemble de la période et des départements s'établit à 1008,1 vols pour 100 000 habitants (voir annexe 7). Ces écarts illustrent clairement les inégalités départementales en matière de criminalité en France

métropolitaine, et justifient une approche économétrique différenciée selon les spécificités locales.

Concernant les variables de contrôle, la majorité d'entre elles présentent une distribution proche de la loi normale, tout en affichant plusieurs valeurs extrêmes (voir Annexes 8 et 9, contenant les histogrammes et boîtes à moustaches). Par exemple, le département des Pyrénées-Orientales (66) enregistre un taux de chômage maximal de 15,325 % en 2015, témoignant d'une situation économique particulièrement tendue à cette période. De même, le département de la Vendée (85) présente le rapport interdécile le plus faible (2,5) sur la période de 2018 à 2021, indiquant une répartition des revenus relativement plus égalitaire que dans les autres départements. Ces variations soulignent l'importance d'intégrer ces variables dans les modèles explicatifs, afin de tenir compte des contextes économiques locaux qui peuvent influencer à la fois l'immigration et la criminalité.

Bien que plusieurs variables présentent des valeurs extrêmes au sein de leur distribution statistique, aucune observation n'a été supprimée. En effet, ces valeurs, bien qu'éloignées de la moyenne, ne constituent pas des anomalies ou des erreurs de saisie, et ne peuvent être considérés comme des valeurs atypiques ou aberrantes. Elles reflètent des spécificités réelles et structurelles de certains départements français. Par exemple, des territoires comme Paris, la Seine-Saint-Denis ou encore les Pyrénées-Orientales se caractérisent par des contextes socio-économiques ou démographiques singuliers, qui se traduisent naturellement par des niveaux atypiques de certaines variables (immigration, criminalité, chômage, etc.). Supprimer ces observations reviendrait à appauvrir l'analyse en effaçant la diversité territoriale, et à biaiser les résultats des modèles économétriques en excluant des cas essentiels à la compréhension des dynamiques étudiées. Dès lors, ces valeurs extrêmes sont conservées volontairement afin de garantir une analyse exhaustive et représentative des disparités départementales en France métropolitaine.

De plus, il est pertinent de s'intéresser aux relations de corrélation entre les différentes variables mobilisées dans cette étude. La matrice de corrélation présentée en annexe 10 permet d'analyser de manière synthétique les interdépendances entre la variable explicative, les variables de contrôle et les différents indicateurs de criminalité. Cette visualisation offre un aperçu global des dynamiques relationnelles au sein du jeu de données, facilitant l'identification de corrélations fortes (positives ou négatives) susceptibles de guider l'interprétation des modèles économétriques.



On observe ainsi que la majorité des variables présentent une forte corrélation positive entre elles. Concernant le taux d'immigration, il apparaît logiquement qu'il est fortement corrélé positivement à la part d'étrangers dans la population totale, et négativement à la part de Français. D'autres corrélations sont aussi plus ou moins explicables, comme la médiane du niveau de vie qui est fortement corrélé positivement avec le taux de profession intermédiaire, le taux de cadre et profession supérieur ou encore le taux de crimes violents.

### **c) Tests de validation économétriques**

Afin de valider plusieurs hypothèses fondamentales en économétrie et d'assurer la robustesse des interprétations causales, différents tests statistiques ont été effectués durant l'analyse.

Tout d'abord, un test de multicolinéarité a été mené sur l'ensemble des variables de contrôles et la variable explicative. Ce test repose sur le calcul du VIF. Un seuil de 10 est en général admis dans la littérature comme étant indicateur d'une multicolinéarité préoccupante. Dans notre cas, tous les VIF sont inférieurs à ce seuil, ce qui indique une faible corrélation linéaire entre les variables et garantit la fiabilité des coefficients estimés (voir annexe 11 pour le graphique du test de multicolinéarité VIF).

Ensuite, l'hypothèse d'homoscédasticité des erreurs, essentielle pour la validité des tests statistiques classiques, a été vérifiée à l'aide du test de Breusch-Pagan.<sup>52</sup> Ce test a été appliqué à chacune des régressions en effets fixes simples, révélant une présence significative d'hétéroscédasticité dans les résidus (voir annexe 12). C'est en réponse à cette violation, que les modèles finaux ont été estimés en utilisant des erreurs standards robustes clusterisées au niveau départemental et annuel.

Enfin, un test de stabilité des variables de contrôle a été réalisé afin d'évaluer la sensibilité des coefficients estimés du taux d'immigration à l'inclusion progressive des contrôles. Cette méthode, illustrée à l'aide d'un *coefplot*, permet de visualiser la stabilité des coefficients à travers plusieurs spécifications. Les résultats montrent que l'ajout des

---

<sup>52</sup> Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.

variables de contrôle n'entraîne quasiment pas de variation substantielle dans les coefficients de la variable d'intérêt. Néanmoins, les variables relatives à la professions intermédiaires et l'âge adultes (30-45 ans) sont celles qui semblent le plus influencer les résultats des coefficients du taux d'immigration. Ces variables sont tout de même conservées dans les estimations finales car elles ne changent pas le signe ni la significativité des estimations. Néanmoins, ce test permet de renforcer la validité des résultats et suggère qu'aucune variable de contrôle ne capte de manière disproportionnée les effets estimés. Le graphique sur l'effet estimé de l'immigration sur le taux de vols sur victimes en annexe 13, représente l'impact de l'inclusion progressive des variables de contrôles. Les points noirs indiquent les valeurs estimées de l'impact du taux d'immigration sur le taux de vols sur victimes, tandis que les barres représentent les intervalles de confiance à 95 %.

Ces différents tests contribuent à la rigueur méthodologique de l'analyse et assurent que les conclusions tirées reposent sur des fondations économétriques solides.

## V - Analyse des résultats

Après avoir défini le cadre théorique, méthodologique et descriptif de l'impact de l'immigration sur les indicateurs de criminalité, nous présentons dans cette partie les résultats empiriques obtenues des différents modèles utilisés et des modèles finaux. Cette section a pour objectif de mettre en évidence les relations statistiques significatives ou non entre la variable explicative et les indicateurs expliqués.

Dans un premier temps, un modèle de régression simple est présenté et critiqué. Ensuite, nous détaillerons les résultats des estimations économétriques finales, en exposant les coefficients significatifs, la capacité des modèles à expliquer la part de la variance intra-départemental, puis en interprétant les signes et les impacts sur la criminalité. Enfin, nous discuterons de l'implication de ces résultats à la lumière de la littérature existante et des limites potentielles de notre analyse. A noter que l'ensemble des modèles intermédiaires effectués ne sont pas analysés, seul un tableau résumant les résultats est présenté.

La régression simple effectuée sur l'ensemble des huit indicateurs de criminalité, visant à estimer une relation linéaire simple entre le taux d'immigration et les taux de criminalité, met en évidence une relation positive et statistiquement significative au seuil de 1 % pour l'ensemble des indicateurs considérés (voir annexe 14 et 15 pour une visualisation des résultats). En d'autres termes, si l'analyse se terminait ci-après, on pourrait conclure souvent à tort, que l'immigration est un facteur explicatif de l'augmentation de la criminalité en France. Cependant, cette interprétation serait hâtive, car ce serait ne pas prendre en compte les spécificités intra-départementales et temporelles. En effet, une régression linéaire met simplement l'ensemble des valeurs observées sur un plan euclidien puis minimise la somme des carrés des écarts entre les valeurs observées et les valeurs prédites. Ce modèle ne permet pas de capturer les dynamiques temporelles ni les hétérogénéités spatiales entre les départements, comme les différences socio-économiques, les politiques locales ou encore les contextes démographiques spécifiques à chaque territoire.

Ainsi, cette approche constitue une indication générale sur l'existence d'une relation entre l'immigration et la criminalité, mais demeure insuffisante pour tirer des conclusions robustes. Il est donc nécessaire d'avoir recours à une méthode économétrique plus

élaborées, intégrant des effets fixes temporels et spatiaux, afin de mieux contrôler les biais potentiels liés aux caractéristiques non observées.

Les modèles finaux intègrent donc des effets fixes départementaux et annuelles, des variables de contrôles socio-économiques et des clusters sur les dépendances temporelles et départementales, afin d'isoler l'impact de l'immigration sur la criminalité de l'ensemble des facteurs qui pourraient biaiser ces résultats.

Pour cette spécification, les résultats des tableaux ci-dessous montrent que pour les indicateurs « *taux crime cheques* », « *taux crime vols\_vehicules* », « *taux crime stupefiants* », « *taux crime vols victime* », « *taux crime violence* » et « *taux crime autres* », le taux d'immigration a un impact statistiquement significatif au seuil de 1 %, 5 % ou 10 %. Tandis que « *taux crime plaintes* » à un impact statistiquement significatif au seuil de 1 % (p-value = 0.00385453) et « *taux crime infractions pub* » à un impact statistiquement significatif au seuil de 10 % (p-value = 0.051980). L'immigration a un impact négatif sur ces deux types de criminalité respectifs (voir tableaux 2 et 3 et annexes 16 à 21).

Plus précisément, toutes choses égales par ailleurs, une augmentation d'un point de pourcentage du taux d'immigration dans un département donné au cours du temps (par exemple passer de 10 % à 11 % d'immigrés dans la population totale) est associée à une baisse de 69,83 plaintes (infractions signalées par des plaignants) pour 100 000 habitants. De plus, toutes choses égales par ailleurs, une augmentation d'un point de pourcentage du taux d'immigration dans un département au cours du temps est associée à une diminution de 40,44 infractions contre des institutions ou des établissements publics pour 100 000 habitants. D'autre part, pour l'ensemble des autres indicateurs de criminalités mesurés, l'immigration n'a aucun effet sur leurs variations.

Il est important de noter que la variable *taux d'immigration* est exprimé en proportion dans le modèle (par exemple, 0.03 signifie 3 % d'immigrés dans la population). Par conséquent, le coefficient estimé de - 6983,30 représente l'effet d'une hausse de 1 unité, soit un passage théorique de 0 % à 100 % d'immigrés, ce qui n'est évidemment pas réaliste. Pour interpréter correctement ce coefficient, il faut raisonner en point de pourcentage, c'est-à-dire qu'une augmentation de 1 point de pourcentage du taux d'immigration (par exemple, de 3 % à 4 %, soit une variation de 0.01 en proportion) est donc associée, toutes choses égales par ailleurs, à une baisse d'environ 69,83 plaintes (-

6983,30  $\times$  0,01) pour 100 000 habitants. Cette précision est essentielle pour éviter toute surestimation des interprétations.

Ces diminutions importantes suggèrent qu'une plus forte proportion d'immigrés est corrélée à une réduction nette de la criminalité enregistrée par plaintes et par infractions contre des institutions ou des établissements publics. Ces résultats, ne reflète donc pas une comparaison entre départements, mais bien l'évolution de la criminalité en fonction des changements du taux d'immigration dans un même territoire au fil des années, toutes choses égales par ailleurs.

En ce qui concerne la qualité explicative des modèles, des  $R^2$  ajustée d'environ 0,89 pour le taux de crimes enregistré par des plaignants et d'environ 0,91 pour le taux d'infraction sur des institutions ou des établissements publics, indiquent que les modèles expliquent près de 90 % de la variance totale des taux de criminalités. Ce qui reflètent une excellente capacité explicative globale, bien qu'il faille rappeler que ces  $R^2$  sont surestimé par les effets constants des effets fixes.

De surcroît, les *Within  $R^2$*  (ou  $R^2$  « interne »), fournissent plus d'informations dans une régression à effets fixes. Le *Within  $R^2$*  d'environ 0.40 pour le taux de plaintes signifie que 40 % des variations annuelles au sein d'un même département sont expliquées par les variables du modèle. Le *Within  $R^2$*  d'environ 0.20 pour le taux d'infraction sur des institutions ou des établissements publics signifie également que 20 % des variations annuelles au sein d'un même département sont expliquées par les variables du modèle. Ces valeurs modérées des *Within  $R^2$*  reflètent soit les limites des variables explicatives retenues pour modéliser la criminalité, soit l'instabilité inhérente à certains phénomènes sociaux. Ces deux indicateurs ne doivent toutefois pas être interprétés comme des indicateurs de la validité des estimations de l'effet de l'immigration sur les taux de criminalité.

Pour ce qui est des contrôles, certaines de ces variables apparaissent de manière récurrente comme statistiquement significatives dans l'ensemble des modèles estimés, ce qui témoigne la stabilité de leur effet. Par exemple, la part des jeunes adultes de 15 à 29 ans a en général un effet positif et significatif sur les différents indicateurs. La part des personnes en professions intermédiaires est majoritairement significative et négative. Enfin, le rapport interdécile est en général responsable d'une augmentation légère de la

criminalité de façon significative, ce qui suggère qu'une plus grande inégalité économique peut contribuer à une hausse des actes délictueux, en cohérence avec certaines théories de la criminalité portée sur les tensions socio-économiques.

Tableau 2 : Résultats de la régression finale le *taux crime plaintes*

```
===== Résultats pour : taux_crime_plaintes =====
OLS estimation, Dep. Var.: taux_crime_plaintes
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate   Std. Error   t value   Pr(>|t|)
TAUX_IMMIGRATION      -6983.302482  2046.159388  -3.412883  0.00385453 **
PART_HOMMES             7366.435291  4524.823487   1.628005  0.12434115
PART_0_14               6879.198459  3176.257948   2.165819  0.04684742 *
PART_15_29             3822.694313  2175.490440   1.757164  0.09927550 .
PART_30_44            -13135.481577  2945.195280  -4.459970  0.00045849 ***
PART_CS1_AGRICULTEURS  -8276.416740  5672.125323  -1.459139  0.16515009
PART_CS4_PROF_INTERMEDIAIRES  5519.143278  3244.146920   1.701262  0.10952459
PART_CS5_EMPLOYES      -973.415189  2520.396522  -0.386215  0.70476040
PART_CS6_OUVRIERS      -76.494959  2580.852589  -0.029639  0.97674540
Taux_chomage           839.490008  1087.048800   0.772265  0.45195420
MEDIANE                 0.029866    0.023037   1.296423  0.21442077
RAPPORT_INTERDECILE     45.357145   16.430503   2.760545  0.01457365 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 75.1      Adj. R2: 0.893878
                Within R2: 0.401957
```

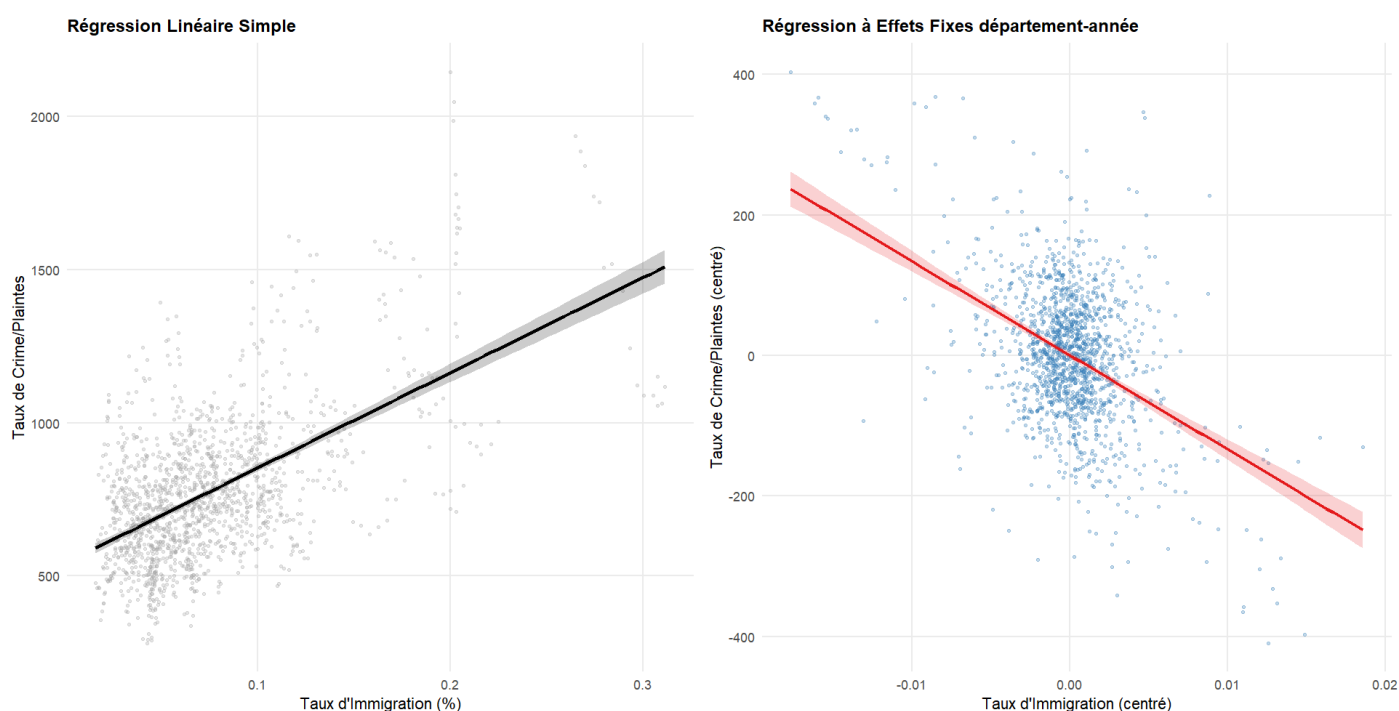
Tableau 3 : Résultats de la régression finale pour le *taux crime infraction publique*

```
===== Résultats pour : taux_crime_infractions_pub =====
OLS estimation, Dep. Var.: taux_crime_infractions_pub
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate   Std. Error   t value   Pr(>|t|)
TAUX_IMMIGRATION      -4043.73768  1915.668902  -2.110875  0.051980 .
PART_HOMMES            -4732.57403  4931.766279  -0.959610  0.352465
PART_0_14              164.62123  3018.422727   0.054539  0.957226
PART_15_29            9138.14062  3237.440244   2.822644  0.012858 *
PART_30_44             69.71698  3426.270169   0.020348  0.984034
PART_45_59            -7507.95861  2961.538281  -2.535155  0.022861 *
PART_CS2_ARTISANS_COMMERCANTS -7427.45416  4750.712936  -1.563440  0.138797
PART_CS4_PROF_INTERMEDIAIRES  1653.54006  2567.800951   0.643952  0.529331
PART_CS5_EMPLOYES     -5163.71303  2323.331478  -2.222547  0.042044 *
Taux_chomage          -1204.15046  958.367757  -1.256460  0.228165
MEDIANE                 0.04862    0.020828   2.334340  0.033894 *
RAPPORT_INTERDECILE     46.09136   17.423553   2.645348  0.018361 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 73.9      Adj. R2: 0.913441
                Within R2: 0.202342
```

Les différentes spécifications économétriques, à savoir la régression simple, et les régressions à effets fixes avec erreurs clusterisées et intégration de variables de contrôles aboutissent à des résultats diamétralement opposés, mais reflètent surtout des niveaux de robustesse très contrastés. La régression simple bien que révélatrice d'une corrélation apparente, ne permet aucune interprétation causale en raison de sa faible capacité à contrôler les biais potentiels. À l'inverse, les modèles finaux construits avec une méthodologie plus rigoureuse permettent une analyse plus fiable et une interprétation plus crédible d'un lien de causalité existant entre l'immigration et la criminalité. La figure 4 illustre clairement les écarts entre ces approches, passant d'une corrélation positive à une corrélation négative entre le taux d'immigration et le taux de plaintes, ce qui souligne l'importance du choix méthodologique dans l'inférence économétrique.

Figure 4 : Comparaison visuelle d'une régression linéaire simple et du modèle final entre le taux d'immigration et le taux de plaintes



Pour conclure, les résultats des modèles finaux conduisent en majorité à l'absence de lien statistiquement significatif entre le taux d'immigration et la plupart des indicateurs de criminalité. Toutefois, deux exceptions notables émergent : le taux de plaintes et le taux d'infractions commises à l'encontre des institutions ou établissements publics. Pour ces deux catégories, le taux d'immigration présente un effet négatif et significatif, suggérant qu'une hausse de l'immigration est associée à une diminution de ces types de criminalité.

Ces relations peuvent s'expliquer de diverses façons, d'abord il se peut qu'il y ait un effet d'auto-sélection des immigrés, étant donné qu'ils ne constituent pas un groupe entièrement aléatoire, mais sont souvent des personnes ayant fait preuve de motivation, d'effort et de résilience pour migrer. Ce phénomène d'auto-sélection peut réduire la probabilité de comportement délictueux, notamment pour les délits graves ou violents. Ensuite, cela peut se traduire par une meilleure intégration et insertion sociale, en effet, la mise en place de réseaux communautaires, de politiques locales d'intégration efficaces ou encore de structures économiques locale favorable à l'immigration peut réduire les incitations à commettre des crimes et délits. Enfin, les infractions enregistrées par les plaignants dépendent uniquement du nombre de signalement par les victimes. Or, dans des zones où la population est plus diverse, il se peut que certaines victimes ne signalent pas ou que les services de police se focalisent sur d'autres priorités, affectant les taux enregistrés.

De manière générale, ces résultats sont cohérents avec une large partie de la littérature scientifique qui remettent en question l'existence d'un lien direct entre immigration et insécurité. Ils contribuent ainsi à déconstruire certains préjugés tenaces selon lesquels l'immigration favoriserait la criminalité. Au contraire, ils mettent en lumière la nécessité d'un débat public mieux informé, fondé sur des données empiriques rigoureuses plutôt que sur des perceptions erronées.



## VI - Analyse des estimations des bases alternatives

Cette partie se focalise sur les résultats des estimations dans le cas de scénarios alternatifs, où les bases de données utilisées sont construites différemment et nécessitent un regard approfondi sur leurs estimations. L'objectif est d'apporter un regard critique et approfondi sur la robustesse des estimations, afin d'évaluer dans quelle mesure les résultats généraux de l'étude se maintiennent. Dans un premier temps, nous présentons les résultats issus de la base du recensement harmonisé (*Saphir*), avant d'analyser dans un second temps, ceux de la base de données des indicateurs de criminalité revisitée depuis 2016, dont les spécificités méthodologiques conduisent à des résultats légèrement divergents.

Dans cette section, l'analyse détaillée de la méthodologie appliquée n'est pas reprise, en raison de sa similarité avec celle effectuée précédemment. De même, la présentation des modèles économétriques est volontairement omise, puisqu'il s'agit des mêmes modèles finaux que ceux utilisés dans l'analyse principale. Il est néanmoins important de souligner que l'ensemble des démarches économétriques et tests de validité menés dans la section précédente ont également été réalisés sur les bases de données alternatives présentées ici. Ces éléments ne sont pas reproduits afin d'éviter les redondances et de préserver la lisibilité de l'analyse.

### 1) Base de données du recensement harmonisé (*Saphir*)

La base *Saphir*, est une base de données historique harmonisée des recensements de la population couvrant la période de 1968 à 2021 et conçu par L'INSEE. Elle intègre les données issues des recensements quinquennaux de la population en France métropolitaine et dans les DOM (depuis 1982), et a pour objectif de permettre des comparaisons temporelles fiables et enrichis malgré les évolutions des concepts et des mesures statistiques. La base *Saphir* propose une quarantaine de variables sociodémographiques harmonisées, portant notamment sur l'âge, le sexe, la nationalité, la catégorie socioprofessionnelle, le niveau de diplôme, la situation conjugale, le type d'activité, ainsi que des informations sur la localisation de résidence, de naissance et de travail. Chaque valeur correspond à un individu décrit par ses caractéristiques

individuelles et de ménages. Cette harmonisation de la base *Saphir* constitue une ressource précieuse pour les analyses longitudinales, en permettant l'étude cohérente de phénomènes sociaux à travers plusieurs recensements et à partir d'une source unique.<sup>53</sup>

Toutefois, l'utilisation de cette base présente certaines limites dans le cadre de cette étude. En particulier, le croisement avec la base des indicateurs de criminalité ne permet d'observer que quatre années communes (1999, 2010, 2015 et 2021). Ainsi, l'analyse de l'impact de l'évolution du taux d'immigration sur les différents indicateurs de criminalité est restreinte à ces quatre années d'observation espacés dans le temps. Par ailleurs, la base *Saphir* ne contient pas d'informations sur les niveaux de richesse par département, en raison du manque de données disponibles. Pour en savoir plus sur la construction de cette base, voir annexe 25, page 72, qui explique plus en détail les démarches effectuées pour construire les différentes variables.

Le tableau 4 ci-dessous présente les résultats des modèles finaux pour la base *Saphir*.

Tableau 4 : Résultats des régressions finale en effets fixes avec variables de contrôles et clusters pour la base *Saphir*

Effet du taux d'immigration selon la variable dépendante

Variable dépendante	Coefficient	Erreur standard	P-value	Signif.
taux_crime_plaintes	-4112.953	1266.912	0.048	*
taux_crime_vols_vehicules	3884.276	3818.395	0.384	
taux_crime_stupefiants	1356.625	734.642	0.162	
taux_crime_vols_victime	-1105.314	1410.055	0.490	
taux_crime_violences	-466.594	657.465	0.529	
taux_crime_autres	3298.805	1369.665	0.095	.
taux_crime_infractions_pub	109.919	1885.537	0.957	

<sup>53</sup> RP Historiques et Données harmonisées des recensements de la population (Saphir) | Insee

Le tableau montre que pour les indicateurs « *taux crime cheques* », « *taux crime vols vehicules* », « *taux crime stupefiants* », « *taux crime vols victime* », « *taux crime violences* » et « *taux crime infractions pub* », le taux d'immigration n'a pas d'impact statistiquement significatif au seuil de 1 %, 5 % ou 10 %. Autrement dit, il n'existe pas de relation significative entre le taux d'immigration et ces différents indicateurs.

De plus, le taux d'immigration a un impact statistiquement significatif sur les indicateurs « *taux crime plaintes* » et « *taux crime autres* », aux seuils respectifs de 5 % (p-value  $\approx 0.048$ ) et 10 % (p-value  $\approx 0.095$ ). Concernant, l'indicateur de criminalité « *taux crime plaintes* », l'immigration a un impact négatif sur celui-ci. En effet, une augmentation d'un point de pourcentage du taux d'immigration dans un département donné au cours du temps, est associée à une baisse de 41,13 plaintes (infractions signalées par des plaignants) pour 100 000 habitants.

En revanche, l'immigration a un impact statistiquement positif sur l'indicateur de criminalité « *taux crime autres* ». De fait, une augmentation d'un point de pourcentage du taux d'immigration, est associée à une hausse de 32,99 crimes « *autres* » (infractions diverses enregistrées comme "procédures").

Les résultats obtenus à partir de la base alternative *Saphir* sont globalement en adéquation avec ceux issus de la base initiale et avec la littérature. Cette stabilité des coefficients estimés suggère que les relations mises en évidence ne sont pas uniquement dues à des spécificités propres à une seule source de données. En particulier, l'effet estimé de l'immigration sur le taux de plaintes confirme l'existence d'un lien négatif récurrent entre immigration et délinquance déclarée par les victimes.

D'un autre côté, les résultats relatifs à l'impact de l'immigration sur le taux d'infractions diverses (correspondant aux procédures administratives ou judiciaires enregistrées par les forces de l'ordre) appellent à une interprétation plus prudente. Contrairement aux autres types de criminalité, cette catégorie regroupe un ensemble plus hétérogène d'actes délictueux susceptibles d'être influencés par des pratiques policières, des modifications réglementaires ou des effets de ciblage administratif, plutôt que par une dynamique criminelle proprement dite. De plus, la nature intrinsèque de certaines infractions, qui sont directement liées à l'immigration, favorisent la relation significative entre

l'immigration et cet indicateur de criminalité. En effet, parmi les infractions diverses on y retrouve les *Autres infractions à la police des étrangers*, le *Travail clandestin*, les *Faux documents d'identité*, les *Délits interdiction de séjour et de paraître*, ou encore les *Infractions aux conditions générales d'entrée et de séjour des étrangers*, qui sont tous des délits relatifs aux conditions des étrangers (dont la majorité sont des immigrés). Dès lors, le lien statistique observé entre immigration et cette catégorie de crimes pourrait refléter des biais institutionnels de déclaration ou de traitement. Cependant, l'estimations de cet effet reste intéressant et nécessite une considération dans l'analyse des résultats.

## **2) Base de données des indicateurs de criminalité revisité depuis 2016**

La base de données alternative des indicateurs de criminalité revisité depuis 2016, est une base mise à disposition par le SSMSI, et fournit des données annuelles détaillées sur les crimes et délits enregistrés par la police et la gendarmerie nationales. Ces données sont disponibles aux niveaux communal, départemental et régional, selon le lieu de commission des faits. Elles contiennent des informations détaillées sur les infractions constatées et regroupe des catégories statistiques plus fines sur la criminalité que celles de « l'état 4001 ». Depuis 2016, l'enrichissement progressif de cette base permet d'analyser les violences intrafamiliales ou l'usage de stupéfiants, et de suivre l'évolution géographique de la délinquance sur la base d'indicateurs fiabilisés. La base exclut les infractions trop peu fréquentes, notamment dans les petites communes, et respecte le secret statistique. Elle est également régulièrement actualisée pour y intégrer de nouveaux indicateurs et affiner les analyses territoriales, offrant ainsi un outil précieux pour l'étude de l'insécurité sur le territoire français.<sup>54</sup>

Cette base couvre 18 indicateurs principaux, dont les homicides, violences sexuelles, cambriolages, vols de véhicules, vols avec ou sans violence, destructions volontaires, usage et trafic de stupéfiants, etc. Cette base permet aussi la construction d'un calculs

---

<sup>54</sup> Bases statistiques communale, départementale et régionale de la délinquance enregistrée par la police et la gendarmerie nationales (base) - [data.gouv.fr](http://data.gouv.fr)

standardisés par logements pour les taux de cambriolages, en supplément de ceux réalisées par population.

Toutefois, les données présentent plusieurs limites qu'il convient de noter. D'abord, des délais peuvent exister entre la commission de l'infraction et son enregistrement, ce qui peut introduire un léger décalage temporel dans les séries. Par ailleurs, les données sont comptabilisées selon une règle stricte (un fait, une localisation, un mis en cause), rendant leur agrégation complexe. Enfin, les infractions sont enregistrées selon leur qualification initiale, sans tenir compte d'éventuelles requalifications judiciaires ultérieures, ce qui peut parfois biaiser certaines catégories de délits.<sup>55</sup>

Les différents indicateurs de criminalité ont toutefois été restructurés afin de ne conserver que sept variables explicatives, dans le but de limiter la surreprésentation des catégories et de simplifier l'analyse. Les regroupements ont été effectué selon une méthodologie rigoureuse et tenant compte de leurs unités de comptes. Les indicateurs regroupés sont présentés dans le tableau 5 ci-dessous.

Tableau 5 : Présentation des indicateurs de criminalités regroupés

Indicateurs	Description	Catégories incluses
<i>taux crimes violent</i>	Violences sexuelles, agressions, homicides	Homicides (y compris coups et blessures volontaires suivis de mort), Tentatives d'homicides, Coups et blessures volontaires sur personnes de 15 ans ou plus (dont violences intrafamiliales, dont autres coups et blessures volontaires), Violences sexuelles
<i>taux crimes escroqueries fraudes</i>	Escroqueries et fraudes aux moyens de paiement	Escroqueries et fraudes aux moyens de paiement
<i>taux crimes vols dégradations</i>	Vols avec ou sans armes et dégradations de bien	Vols avec armes (armes à feu, armes blanches ou par destination), Vols violents sans arme,

<sup>55</sup> Méthodologies de collecte et d'analyse des données pour le suivi et l'évaluation | M&E Guidelines

		Destructions et dégradations volontaires
<i>taux crimes cambriolages</i>	Taux de cambriolages pour 100 000 logements	Cambriolages de logements
<i>taux crimes vols sans violence victimes</i>	Vols sans violence contre des personnes	Vols sans violence contre des personnes
<i>taux crimes vols vehicules</i>	Vols contre des véhicules	Vols de véhicules, Vols dans les véhicules, Vols d'accessoires sur véhicules
<i>taux crimes stupefiants</i>	Crimes liés au stupéfiants	Usage de stupéfiants (dont amendes forfaitaires délictuelles (AFD) pour usage de stupéfiants), Trafic de stupéfiants

Le tableau 6 ci-dessous présente les résultats des modèles finaux pour la base *Criminalité* revisité depuis 2016.

Tableau 6 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour la base *Criminalité* revisité depuis 2016

#### Effet du taux d'immigration selon la variable dépendante

Variable dépendante	Coefficient	Erreur standard	P-value	Signif.
taux_crimes_violent	902.347	2213.639	0.700	
taux_crimes_escroqueries_fraudes	-3986.817	942.443	0.008	**
taux_crimes_vols_degradations	-5018.536	2296.321	0.081	.
taux_crimes_vols_sans_violence_victimes	16913.474	5281.474	0.024	*
taux_crimes_vols_vehicules	408.548	2452.540	0.874	
taux_crimes_stupefiants	336.173	3304.159	0.923	
taux_crimes_cambriolages	-1194.456	3398.842	0.740	

Pour les modalités « *taux crimes violent* », « *taux crimes vols véhicules* », « *taux crimes stupefiants* » et « *taux crimes cambriolages* », le taux d'immigration n'a aucun impact statistiquement significatif au seuil de 1 %, 5 % ou 10 %. Autrement dit, il n'existe pas de relation significative entre le taux d'immigration et ces différents indicateurs.

De surcroît, le taux d'immigration a un impact statistiquement significatif sur les indicateurs « *taux crimes escroqueries fraudes* », « *taux crimes vols dégradations* » et « *taux crimes vols sans violence victimes* », aux seuils respectifs de 1 % (p-value  $\approx 0.008$ ), 10% (p-value  $\approx 0.081$ ) et 5 % (p-value  $\approx 0.024$ ). Concernant, les indicateurs de criminalité « *taux crimes escroqueries fraudes* » et « *taux crimes vols dégradations* », l'immigration a un impact négatif sur eux.

En effet, toutes choses égales par ailleurs, une augmentation d'un point de pourcentage du taux d'immigration dans un département donné au cours du temps, est associée à une baisse d'environ 39,87 escroqueries et fraudes aux moyens de paiement, pour 100 000 habitants. De plus, une augmentation d'un point de pourcentage du taux d'immigration est également associée à une baisse d'environ 50,18 vols avec ou sans armes et dégradations de bien, pour 100 000 habitants.

En revanche, l'immigration a un impact statistiquement positif sur l'indicateur de criminalité « *taux crimes vols sans violence victimes* ». De fait, une augmentation d'un point de pourcentage du taux d'immigration dans un département donné au cours du temps, est associée à une hausse de 169,13 vols sans violence contre des personnes, pour 100 000 habitants.

Les résultats obtenus à partir de la base alternative *Criminalité* sont relativement en adéquation avec ceux issus de la base initiale, ainsi qu'avec une partie de la littérature existante. Par exemple, l'indicateur des *vols avec ou sans armes et dégradations de biens* présente un effet négatif du taux d'immigration, ce qui rejoint les résultats observés précédemment avec le taux d'infractions contre des institutions ou des établissements publics dans la base initiale. Cette relation négative peut s'expliquer par le fait que les populations immigrées tendent à s'installer dans des zones urbaines où la surveillance sociale est plus forte (présence policière accru, réseaux communautaires etc.), ou encore par le rôle dissuasif de la précarité, rendant l'implication dans ce type de criminalité risquée et peu avantageuse.

À l'inverse, l'indicateur du *taux de vols sans violence contre des personnes* affiche un effet significatif positif du taux d'immigration. Ce résultat peut refléter plusieurs dynamiques, d'une part ces délits sont souvent plus répandus dans les zones densément peuplées, marquées par des tensions socio-économiques, et où la concentration d'immigrés est plus forte. D'autre part, certains travaux suggèrent que les facteurs de marginalisation économique peuvent accroître la petite délinquance opportuniste, notamment dans les contextes de chômage élevé ou d'accès limité aux ressources.<sup>5657</sup>

Enfin, l'indicateur du *taux d'escroqueries et fraudes aux moyens de paiement* montre, lui aussi, une relation négative avec le taux d'immigration. Ce résultat est intéressant et peut s'expliquer par le fait que les escroqueries impliquent généralement un niveau d'organisation plus élevé, voire des compétences techniques ou des connaissances administratives précises. Or, les populations immigrées, en particulier récentes, peuvent rencontrer des obstacles linguistiques ou institutionnels qui les rendent moins susceptibles de participer à ce type de criminalité. De plus, leur précarité économique et juridique peut limiter leur implication dans des actes délictueux complexes et fortement pénalisés.

En somme, ces résultats illustrent la nécessité de nuancer l'analyse des liens entre immigration et criminalité, en tenant compte de la nature des infractions et des dynamiques sociales sous-jacentes. Ils soulignent que l'effet de l'immigration n'est ni déterminé ni systématiquement négatif ou positif, mais dépend fortement du type de délit considéré.

---

<sup>56</sup> Paugam, S. (2007). *Les formes élémentaires de la pauvreté*. Presses Universitaires de France.

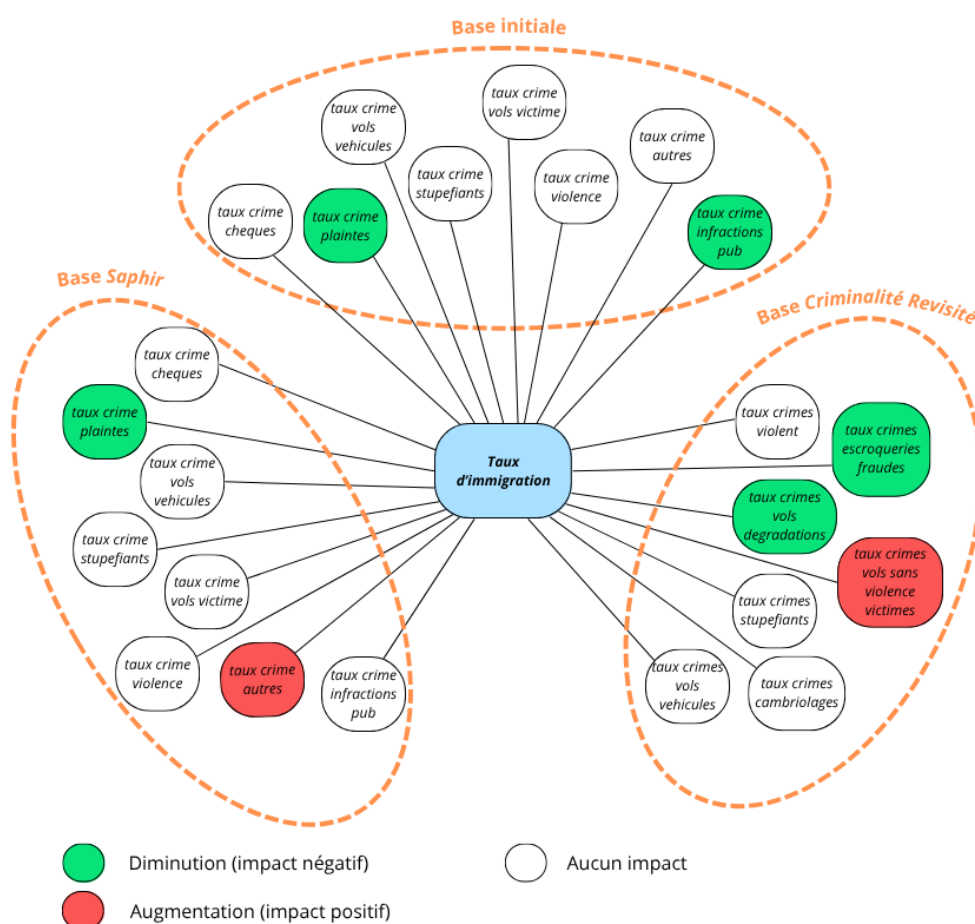
<sup>57</sup> Bourdieu, P. (1979). *La distinction: Critique sociale du jugement*. Les Éditions de Minuit.



## VII – Résumés des résultats finaux

La figure 4 ci-dessous illustre l'effet de l'immigration sur les différents indicateurs de criminalité, en s'appuyant sur les données issues des diverses bases utilisées. Les annexes 22, 23 et 24 présentent de manière détaillée les estimations obtenues selon les différentes spécifications des modèles.

Figure 4 : Graphique de l'impact du taux d'immigration sur les différents indicateurs de criminalité



Source : Canva, Auteur

## VIII - Discussion

Cette section vise à examiner la contribution et les résultats obtenus, et vise à les comparer avec ceux de la littérature existante (a). Nous aborderons ensuite les différentes limites de l'étude (b). Puis nous élaborons des recommandations éclairées suivant les résultats analysés (c).

### a) Analyse des résultats

L'utilisation de multiples techniques de régressions suivi de l'analyse des régressions en effets fixes, accompagnées de variables de contrôle et d'erreurs standard clusterisées, renforce la validité interne de l'étude. Cela permet d'obtenir des résultats fiables et robustes, tout en limitant l'influence des facteurs extérieurs non observés et hétérogènes qui auraient pu biaiser les estimations. Ainsi, les résultats suggèrent avec rigueur l'existence de relations causales entre le taux d'immigration et certains indicateurs de criminalité. Par ailleurs, les différents tests de significativité et de validité statistique, réalisés tout au long de l'analyse, viennent appuyer la qualité des estimations produites.

L'effort méthodologique apporté à la construction et à la sélection des bases de données représente également un atout majeur. L'élaboration de plusieurs bases, issues de différentes sources (dont l'INSEE, le SSMSI ou encore le ministère de l'intérieur), permet de tester la stabilité des résultats selon plusieurs contextes temporels et structurels. Cela contribue à la validité externe de l'étude, rendant les résultats potentiellement généralisables à d'autres contextes ou périodes. Par conséquent, les conclusions obtenues peuvent être mobilisées dans une optique d'aide à la décision publique, dans des situations concrètes au-delà du seul cadre académique de ce mémoire.

L'analyse réalisée sur la base initiale de 2006 à 2021 révèle que l'immigration a un effet significativement négatif sur deux indicateurs de criminalité. Le taux de plaintes : une hausse d'un point de pourcentage du taux d'immigration, dans un département à une année donnée, est associée à une diminution de 69,83 plaintes pour 100 000 habitants. Et, le taux d'infractions contre les institutions ou établissements publics : la même augmentation du taux d'immigration entraîne une baisse de 40,44 infractions pour 100 000 habitants. Pour les six autres indicateurs de criminalité inclus dans cette base, aucun

effet statistiquement significatif n'a été observé. Cela montre que l'immigration n'a pas d'effet systématique sur la criminalité au sens large, mais peut influencer uniquement certains types bien précis.

Les estimations effectuées à partir de la base *Saphir*, couvrant les années 1999, 2010, 2015 et 2021, confirment en partie les résultats précédents. Lorsque le taux d'immigration augmente d'un point de pourcentage, le taux de plaintes diminue de 41,13 pour 100 000 habitants. Cependant, cette base révèle également un effet positif de l'immigration sur les infractions diverses enregistrées comme "procédures", avec une augmentation de 32,99 pour 100 000 habitants. Ces crimes dits « *autres* » englobent des délits diverses administratifs ou infractions hétérogène dont plusieurs d'entre elles sont directement liées à des situations d'immigration.

Avec des indicateurs différents et sur une période de 2016 à 2021, la base *Criminalité* (revisité depuis 2016) confirme certains effets négatifs de l'immigration sur la criminalité. Une hausse du taux d'immigration est liée à une baisse de 39,87 escroqueries et fraudes aux moyens de paiement, pour 100 000 habitants. De même, qu'une hausse du taux d'immigration est associée à une baisse de 50,18 vols avec ou sans armes et dégradations de biens, pour 100 000 habitants. Toutefois, un effet positif ressort également, l'augmentation d'un point de pourcentage du taux d'immigration est liée à une hausse de 169,13 vols sans violence contre des personnes pour 100 000 habitants.

De manière générale, le taux d'immigration ne présente aucun effet significatif sur la majorité des indicateurs de criminalité étudiés. À l'instar de Bell, Machin et Fasani (2013) et de nombreuse étude académique, il n'existe aucun lien entre la variation des crimes violent et des crimes de propriétés, et l'augmentation de l'immigration. Toutefois, lorsqu'un lien statistique existe, il semble être différent selon la nature des infractions. En effet, l'immigration tend à réduire les formes de criminalité nécessitant une certaine organisation ou compréhension du système local comme, les infractions enregistrées par les plaignants, les infractions contre les institutions ou établissements publics, les escroqueries et fraudes aux moyens de paiement ou les vols avec ou sans armes et dégradations de biens. Ainsi, qu'elle est susceptible d'accroître certaines infractions mineures, souvent liées à des contextes de précarité comme les vols sans violence contre des personnes. Ce type de crime peut être liés à des difficultés économiques, à la recherche de ressources ou à une intégration difficile. À l'inverse, porter plainte, frauder un système

bancaire ou s'en prendre à des institutions demande une connaissance des structures locales, qui semble moins accessible aux nouveaux arrivants. Cela suggère que l'intégration sociale et économique joue un rôle central dans les liens observés entre immigration et criminalité.

## **b) Limites**

Malgré ces résultats enrichissants, plusieurs limites intrinsèques méritent d'être soulignées. Dans un premier temps, l'étude se concentre uniquement sur l'immigration régulière, mesurable statistiquement, l'immigration clandestine échappent aux champs d'analyse. De même, elle ne fait pas la distinction selon le profil des immigrés. L'analyse ne distingue pas les immigrés selon leur origine géographique, leur niveau d'étude, leur statut (réfugié, travailleur, étudiant, etc.) ou leur durée de séjour. Or, ces éléments pourraient influencer différemment les dynamiques locales et la criminalité.

Dans un second temps, les estimations ne prennent pas en compte la criminalité non déclarée ou mal classifiée. Par exemple, les violences intrafamiliales ou les délits mineurs, sont souvent sous-déclarés ou mal enregistrés, ce qui peut biaiser les résultats. Ensuite, l'utilisation de données agrégées au niveau départementales ne permet pas de capturer les dynamiques locales plus fines comme les communes ou quartier. Cela limite la capacité de l'étude à saisir des effets potentiellement concentrés dans certains territoires ou zones urbaines sensibles.

De surcroît, bien que les effets fixes permettent de contrôler certaines hétérogénéités non observées, l'étude ne traite pas directement l'endogénéité potentielle entre immigration et criminalité. Une analyse approfondie utilisant les variables instrumentales aurait été pertinente dans ce cadre. Aussi, l'estimation des effets fixes peuvent être légèrement surestimé à la hausse en raison de la présence de valeurs extrêmes qui n'ont pas été exclues de l'analyse. Des observations élevées dans certains départements peuvent artificiellement accroître les écarts moyens associés aux variables muettes, influençant ainsi les résultats.

Par ailleurs, plusieurs variables de contrôles et l'ajout d'enquêtes de victimisations n'ont pu être incluses en raison d'un manque de données disponible à l'échelle départementale.

Le taux d'urbanisation, les motifs migratoires, les politiques d'accueil ou encore les niveaux de santé locaux auraient permis d'affiner les analyses. De plus, les spécificités des DROM n'ont pas été traitées, réduisant la portée de l'analyse sur l'ensemble du territoire français.

Enfin, l'analyse ne prend pas en compte la possibilité d'effets temporels différés. L'impact de l'immigration sur la criminalité pourrait apparaître plusieurs années après l'arrivée des immigrés, ce que les modèles statiques en panel ne permettent pas de capturer pleinement.

### **c) Recommandations**

Au regard de ces résultats, il semble pertinent d'élaborer quelques recommandations visant à améliorer les approches économétriques, ainsi que la qualité de vie de la population. Pour ce faire, il serait pertinent de réaliser des études complémentaires à l'échelle des communes ou des quartiers, afin d'analyser les effets localisés de l'immigration sur certains types de criminalité, en particulier dans les zones urbaines denses, où la concentration d'immigrés est forte et où les dynamiques sociales sont plus complexes.

Il serait judicieux de soutenir des politiques publiques visant à améliorer l'accueil et l'intégration des populations immigrées. Réduire les situations de précarité et faciliter l'insertion économique et sociale permettrait non seulement d'améliorer les conditions de vie de ces populations, mais aussi de réduire certains risques de délinquance et de pauvreté. Ces politiques doivent être différenciées selon les profils des immigrés, car les besoins et les vulnérabilités peuvent fortement varier selon l'âge, l'origine, le niveau de qualification ou le motif de migration. Une approche ciblée permettrait de répondre plus efficacement aux défis d'intégration.

En outre, il est essentiel de lutter contre les stéréotypes et discours stigmatisants, en se basant sur des résultats empiriques solides, plutôt que sur des idées reçues. Une meilleure diffusion des travaux scientifiques rigoureux, notamment à destination des décideurs publics, des médias et de l'opinion générale, est indispensable pour favoriser un débat apaisé et rationnel autour des questions migratoires.

Enfin, il conviendrait également de renforcer la qualité et la précision des données statistiques utilisées dans l'étude des liens entre immigration et criminalité. Il s'agirait notamment de mieux différencier les statuts juridiques des immigrés, d'améliorer la couverture des délits peu signalés ou sous-estimés, et de rendre accessibles des variables socio-économiques plus complètes au niveau territorial.

## Conclusion

Cette étude avait pour objectif d'examiner de manière approfondie la relation causale entre l'immigration et la criminalité, une relation largement explorée à l'échelle internationale mais encore peu étudiée sur le territoire français. Pour ce faire, une analyse en données de panel recourant à des effets fixes, des variables de contrôles socio-économiques et des erreurs standard clusterisées a permis d'établir une base solide, fiable et robuste pour l'identification d'un lien entre immigration et criminalité.

À la suite de la revue de la littérature internationale et nationale couvrant ce sujet, ainsi qu'une présentation détaillée des phénomènes et des facteurs étudiés, une présentation des différentes techniques économétriques mobilisés a été proposée. Les modèles finaux retenus ont ensuite été exposés, suivis d'une analyse rigoureuse de la méthodologie et des tests de validités appliquées. Enfin, une présentation détaillée des résultats obtenus a été accompagnée d'une discussion critique permettant de replacer les conclusions dans leur contexte scientifique et socio-économique.

Les résultats obtenus à l'aide des trois bases de données mobilisés mettent en évidence la prépondérance d'absence de lien statistiquement significatif entre l'immigration et la majorité des indicateurs de criminalité. En particulier, l'absence de lien systématique entre l'immigration et les crimes jugés violents. Néanmoins, certaine relation reste à interpréter avec prudence puisque l'existence de liens entre immigration et criminalité est différente selon la nature des crimes et délits étudiés. L'immigration a en générale un impact négatif sur les crimes et délits qui requièrent une connaissance approfondie du système, des institutions et de la langues locales comme, les *infractions enregistrées par les plaignants*, les *infractions contre les institutions ou établissements publics*, les *escroqueries et fraudes aux moyens de paiement* ou les *vols avec ou sans armes et dégradation de biens*. En outre, l'immigration a un impact positif sur les crimes et délits

non violent relatif à la précarité comme, les *vols sans violence contre des personnes*. Mais aussi, sur des crimes et délits de nature hétérogène comme les *infractions diverses enregistrées comme "procédures"*, où le lien causal est difficile à identifier au sein de cet indicateur.

Finalement, les résultats de cette étude apportent des constats intéressants sur l'existence d'une relation entre ces deux phénomènes. Notamment en mettant en avant les indicateurs de criminalités suspectés d'être impactés par l'augmentation des taux d'immigration au niveau départemental. D'autre part, cette analyse permet de formuler des recommandations sur la base d'une démarche rigoureuse et fiable auprès des décideurs publics et souligne également l'importance de s'appuyer sur ces démarches empiriques avant de tirer des conclusions hâtives sur des phénomènes aussi complexes.

Dans cette optique, il apparaît essentiel de renforcer la sensibilisation du grand public et des institutions à la valeur des études scientifiques fondées sur des données fiables et transparentes. Cela implique notamment de promouvoir des politiques publiques inclusives et progressistes, axées sur l'intégration socio-économique des populations immigrées, et de lutter activement contre les discours stigmatisants et les préjugés infondés.

Enfin, il serait également intéressant de se pencher sur l'impact de l'immigration sur le marché de l'emploi et sur le coût économique des immigrés en France. Mais également d'évaluer l'impact réel des politiques d'accueil, d'accès à l'emploi, au logement ou à l'éducation à destination des immigrés. Cela permettrait d'identifier les dispositifs les plus efficaces, et d'orienter les politiques publiques vers des leviers ayant démontré leur efficacité. Une coopération renforcée entre les acteurs locaux, les collectivités, les associations et les services de l'État, serait aussi un atout pour assurer la mise en œuvre cohérente de ces politiques sur le territoire.

En outre, compte tenu du climat social actuel, marqué par une montée des tensions identitaires et des phénomènes de stigmatisation, il serait pertinent d'élargir le champ de cette recherche à l'ensemble des déterminants influençant la criminalité en France, au-delà du seul facteur migratoire.

# Bibliographie, Sitographie, liens des bases de données

## Bibliographie

- Aoki, Y., & Todo, Y. (2009). Are immigrants more likely to commit crimes? *Evidence from France. Applied Economics Letters*, 16(15), 1537–1541.
- Arellano, M. (1987). Practitioners' corner: Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, 49(4), 445–450.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Alonso-Borrego, C., Garoupa, N., & Vázquez, P. (2012). Does immigration cause crime? Evidence from Spain. *American Law and Economics Review*, 14(1), 165–191.
- Bell, B., Fasani, F., & Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, 95(4), 1278–1290.
- Bianchi, M., Buonanno, P., & Pinotti, P. (2012). Do immigrants cause crime? *Journal of the European Economic Association*, 10(6), 1318–1347.
- Bourdieu, P. (1979). *La distinction: Critique sociale du jugement*. Les Éditions de Minuit.
- Butcher, K. F., & Piehl, A. M. (1998). Cross-city evidence on the relationship between immigration and crime. *Journal of Policy Analysis and Management*, 17(3), 457–493.
- Chalfin, A., & McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1), 5–48.
- Chytilova, H. (2015). Internal and external validity in experimental economics.
- Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica*, 51(2), 277–304.
- Fajnzylber, P., Lederman, D., & Loayza, N. (2002). Inequality and violent crime. *Journal of Law and Economics*, 45(1), 1–39.
- Farrington, D. P. (1986). Age and crime. *Crime and Justice*, 7, 189–250.
- Fougère, D., Kramarz, F., & Pouget, J. (2005). L'analyse économétrique de la délinquance : Une synthèse de résultats récents. *Revue française d'économie*, 19(3), 3–42.
- Fougère, D., Pouget, J., & Kramarz, F. (2009). Youth Unemployment and Crime in France. *Journal of the European Economic Association*, 7(5), 909–938.
- Freeman, R. (1999). The economics of crime. *Handbook of Labor Economics*, 3, 3529–3571.
- Gehrsitz, M., & Ungerer, M. (2017). Jobs, crime, and votes: A short-run evaluation of the refugee crisis in Germany. *ZEW Discussion Paper*, 17-016.
- Grogger, J. (1998). Market wages and youth crime. *Journal of Labor Economics*, 16(4), 756–791.



- Jobard, F., Lévy, R., Lamberth, J., & Névanen, S. (2012). Mesurer les discriminations selon l'apparence : une analyse des contrôles d'identité à Paris. *Revue française de sociologie*, 53(3), 423-451.
- Lagrange, H. (2001). Crime et conjoncture socio-économique. *Revue française de sociologie*, 42(1), 57-79.
- Lechner, M. (2007). A note on endogenous control variables in causal studies. *SIW*.
- Odabasi, S. (2021). The economics of crime and immigration: A panel data analysis. *Pamukkale University Journal of Social Sciences Institute*.
- Paugam, S. (2007). *Les formes élémentaires de la pauvreté*. Presses Universitaires de France.
- Pons, S. (2017). Les crimes ont-ils un genre ? Étude statistique comparée de la criminalité masculine et féminine en Haute-Garonne au XIXe siècle. *Les Cahiers de Framespa*, 25.
- Raphael, S., & Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *Journal of Law and Economics*, 44(1), 259-283.
- Rumbaut, R. G., & Ewing, W. A. (2007). *The myth of immigrant criminality and the paradox of assimilation*. Immigration Policy Center.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.
- Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning.

## Sitographie

- [Accueil - Insee - Institut national de la statistique et des études économiques](#)
- [Bienvenue sur le site du ministère de l'Intérieur](#)
- <https://www.data.gouv.fr/fr/>
- <https://www.insee.fr/fr/accueil>
- [Accueil - data.gouv.fr](#)
- [Immigration : Le débat « peut être apaisé et dépolitisé par la connaissance et le dialogue »](#)
- [Elections européennes : comment la question migratoire a pesé sur le vote des Français](#)
- [Immigration et délinquance : attention aux interprétations trompeuses des chiffres du ministère de l'Intérieur – DE FACTO – Des clés pour mieux s'informer](#)
- [L'essentiel sur... les immigrés et les étrangers | Insee](#)
- [Insécurité et délinquance en 2023 : bilan statistique et atlas départemental de la délinquance enregistrée | Ministère de l'Intérieur](#)
- [Définition - Immigré | Insee](#)
- [Documentation complémentaire sur le recensement | Insee](#)
- [Conseils pour l'utilisation des résultats statistiques | Insee](#)
- [Définition - IRIS | Insee](#)

- [Définition - Criminalité | Insee](#)
- [Insécurité et délinquance en 2023 : bilan statistique et atlas départemental de la délinquance enregistrée | Ministère de l'Intérieur](#)
- [Indicateurs annuels de la victimation et du sentiment d'insécurité issus des enquêtes Cadre de Vie et Sécurité \(CVS\) - data.gouv.fr](#)
- [Endogénéité - SLPV analytics](#)
- [L'essentiel sur... le chômage | Insee](#)
- [Indicateur d'inégalité : le rapport interdécile - La finance pour tous](#)
- [Les migrations des jeunes Européens après la crise de 2008 - Séminaires et colloques de l'Ined - Rencontres scientifiques - Ined - Institut national d'études démographiques](#)
- [RP Historiques et Données harmonisées des recensements de la population \(Saphir\) | Insee](#)
- [Méthodologies de collecte et d'analyse des données pour le suivi et l'évaluation | M&E Guidelines](#)

## Liens des bases de données

### Base initiale (2006-2021)

- [Chiffres départementaux mensuels relatifs aux crimes et délits enregistrés par les services de police et de gendarmerie depuis janvier 1996 - data.gouv.fr](#)
- [Population en 2006 | Insee](#)
- [Population en 2007 | Insee](#)
- [Population en 2008 | Insee](#)
- [Population en 2009 | Insee](#)
- [Population en 2010 | Insee](#)
- [Population en 2011 | Insee](#)
- [Population en 2012 | Insee](#)
- [Population en 2013 | Insee](#)
- [Population en 2014 | Insee](#)
- [Population en 2015 | Insee](#)
- [Population en 2016 | Insee](#)
- [Population en 2017 | Insee](#)
- [Population en 2018 | Insee](#)
- [Population en 2019 | Insee](#)
- [Population en 2020 | Insee](#)
- [Population en 2021 | Insee](#)

### Base de données Saphir (1999, 2010, 2015, 2021)

- [Données harmonisées des recensements de la population 1968-2021 – Données harmonisées des recensements de la population à partir de 1968 | Insee](#)

### Base de données *Criminalité* revisité (2016-2021)

- [Bases statistiques communale, départementale et régionale de la délinquance enregistrée par la police et la gendarmerie nationales \(base\) - data.gouv.fr](#)

## Variables de contrôles

- [Inégalités de revenus et niveau de vie - Indicateurs de structure et de distribution des revenus de 2006 à 2011 | Insee](#)
- [Revenus et pauvreté des ménages en 2012 | Insee](#)
- [Revenus et pauvreté des ménages en 2013 | Insee](#)
- [Revenus et pauvreté des ménages en 2014 | Insee](#)
- [Revenus et pauvreté des ménages en 2015 | Insee](#)
- [Revenus et pauvreté des ménages en 2016 | Insee](#)
- [Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local – Principaux résultats sur les revenus et la pauvreté des ménages en 2017 | Insee](#)
- [Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local – Principaux résultats sur les revenus et la pauvreté des ménages en 2018 | Insee](#)
- [Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local en 2019 – Principaux résultats sur les revenus et la pauvreté des ménages en 2019 | Insee](#)
- [Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local en 2020 – Principaux résultats sur les revenus et la pauvreté des ménages en 2020 | Insee](#)
- [Principaux indicateurs sur les revenus et la pauvreté aux niveaux national et local en 2021 – Principaux résultats sur les revenus et la pauvreté des ménages en 2021 | Insee](#)
- [Police municipale : Effectifs par commune - data.gouv.fr](#)
- [Séries Taux de chômage localisé | Insee](#)

## Packages Python et R

Packages R : arrow, dplyr, tidyr, readxl, ggplot2, sf, tidyverse, gridExtra, writexl, plm, fixest, modelsummary, broom, purr, knitr, car, FactoMineR, factoextra, scales, lmtest, patchwork, officer, flextable, kableExtra

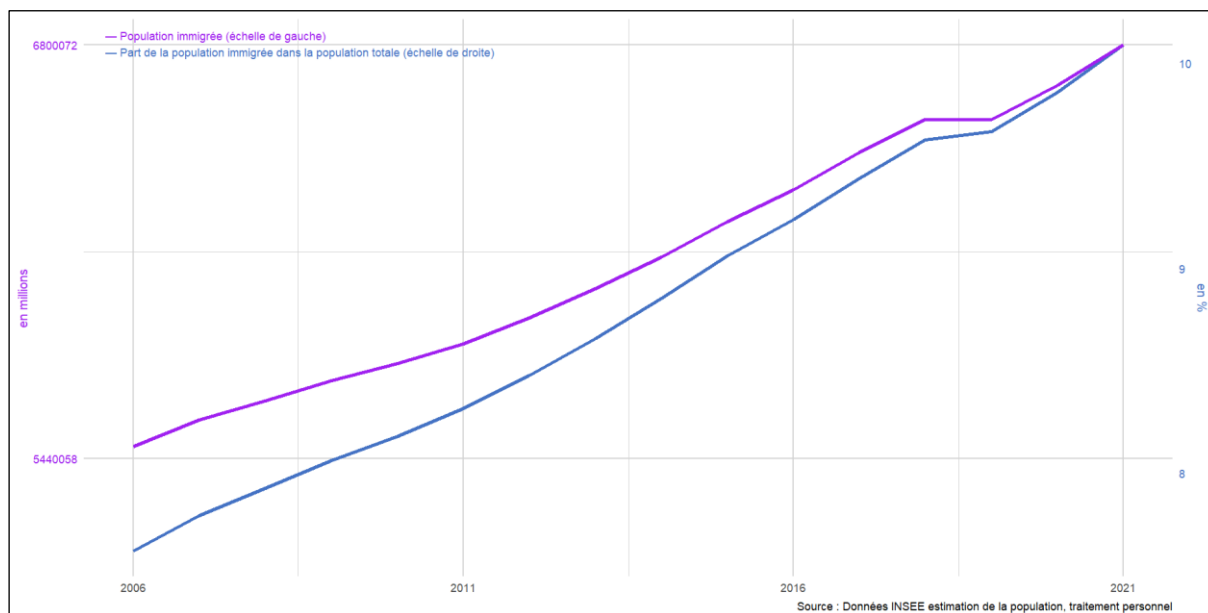
Packages Python : Panda, Re

## Cours Master ECAP

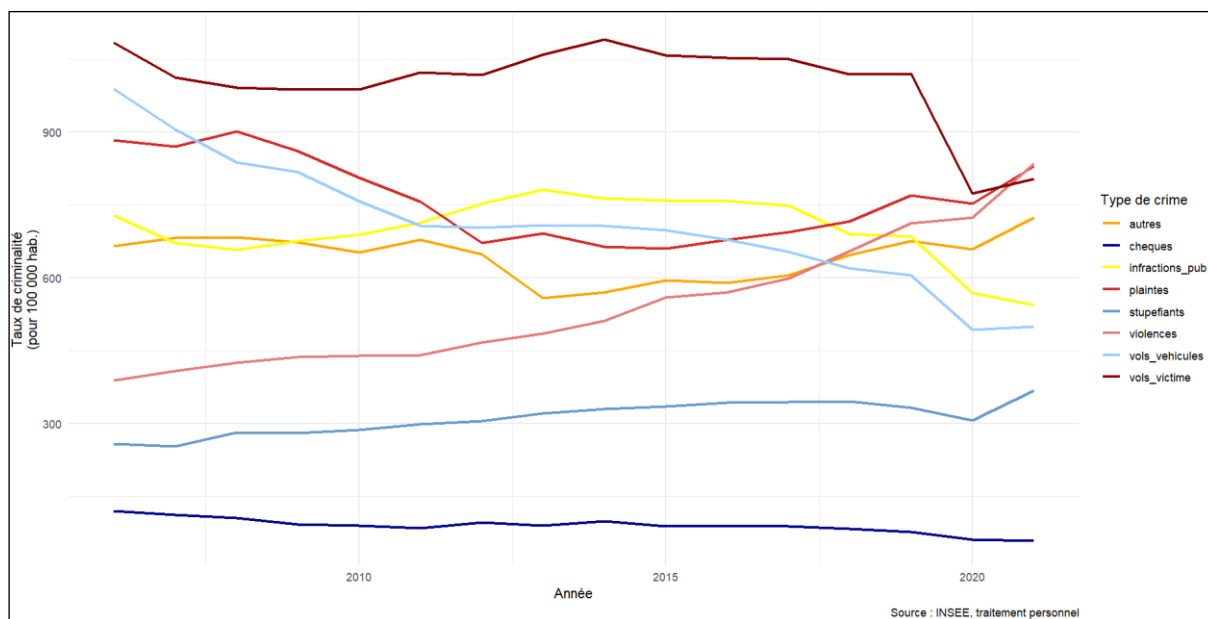
*Econométrie linéaire avancé, Python avancé, Public Policy Evaluation II et R avancé*

## ANNEXES

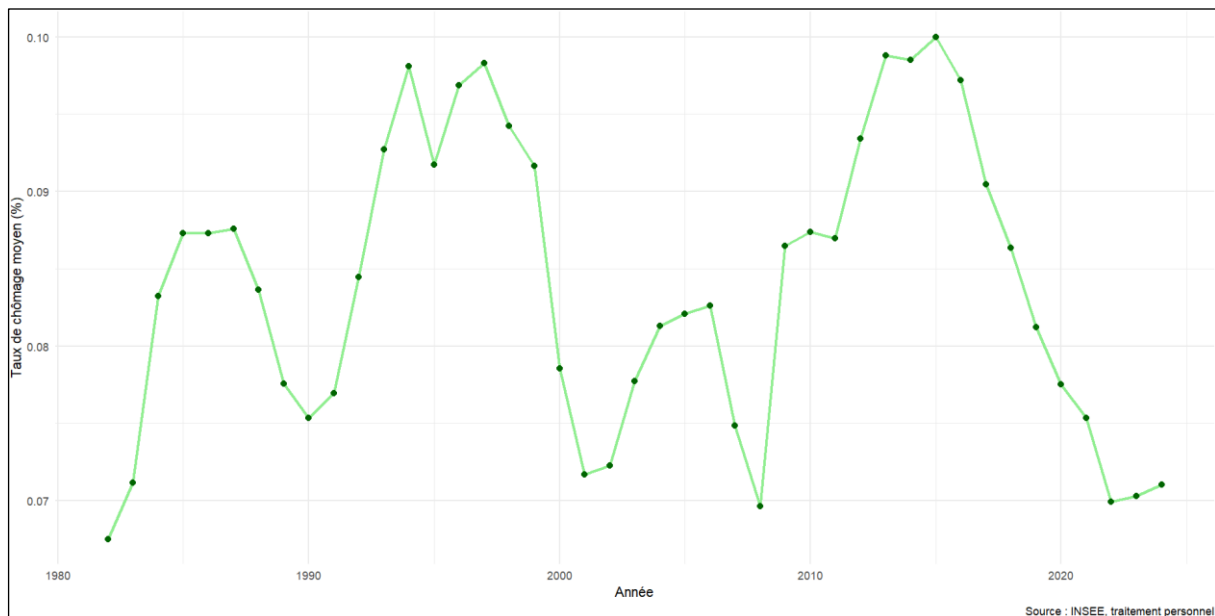
### Annexe 1 : Évolution du taux d'immigration et du nombre d'immigrés de 2006 à 2021 en France



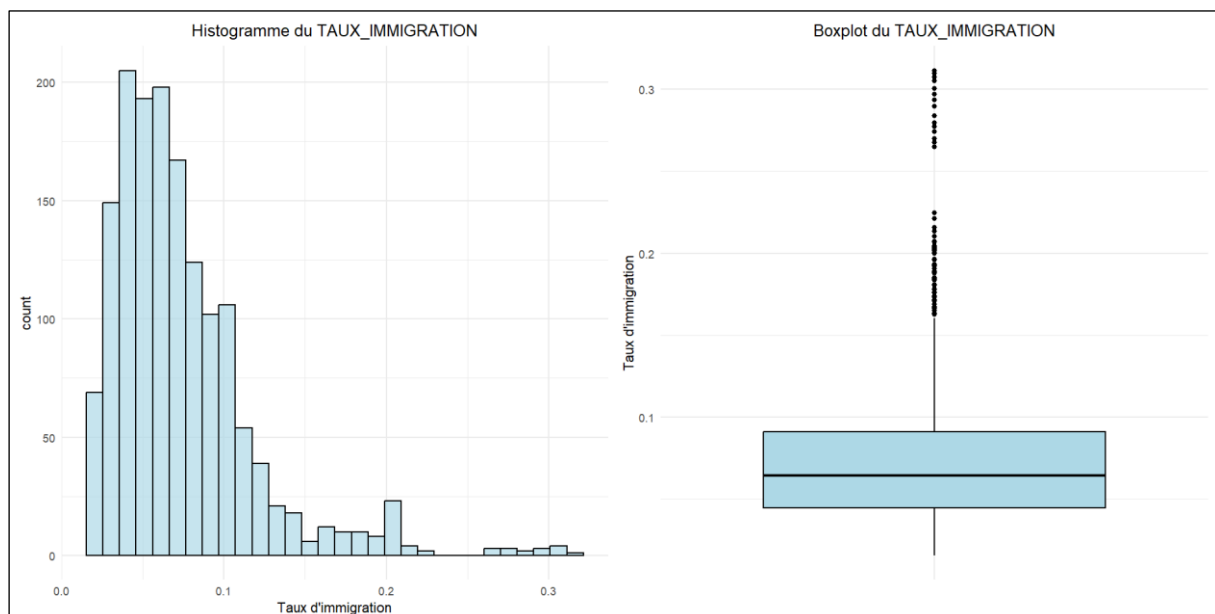
### Annexe 2 : Évolution des indicateurs du taux de criminalité par années de 2006 à 2021



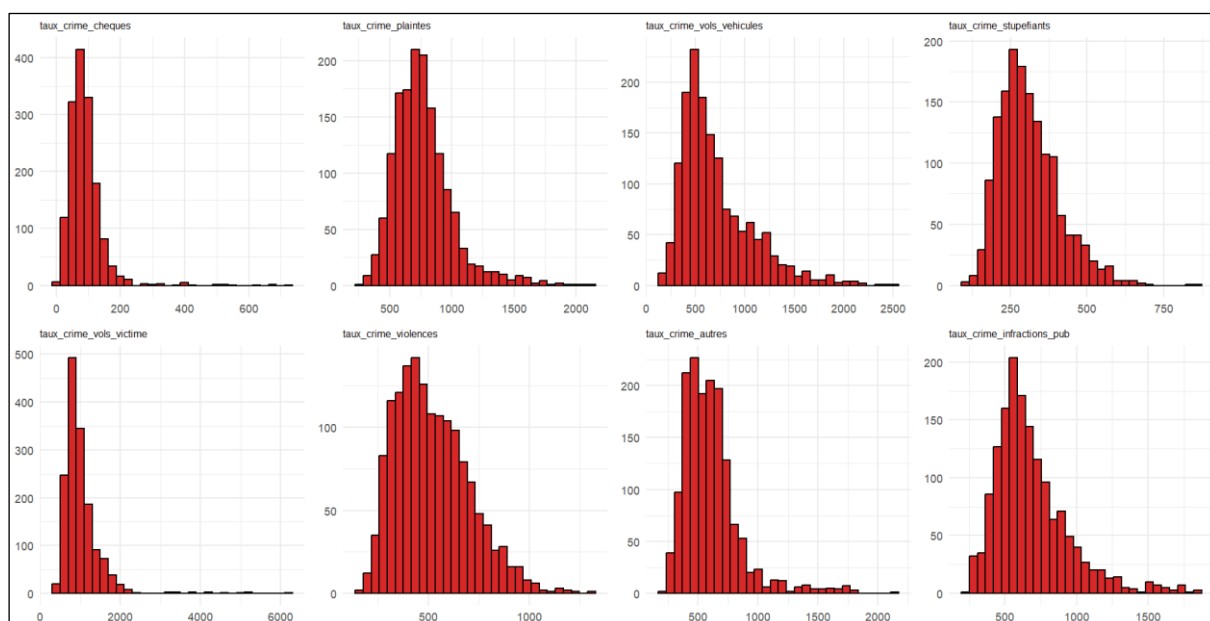
### Annexe 3 : Évolution du taux de chômage moyen en France Métropolitaine de 1982 à 2024



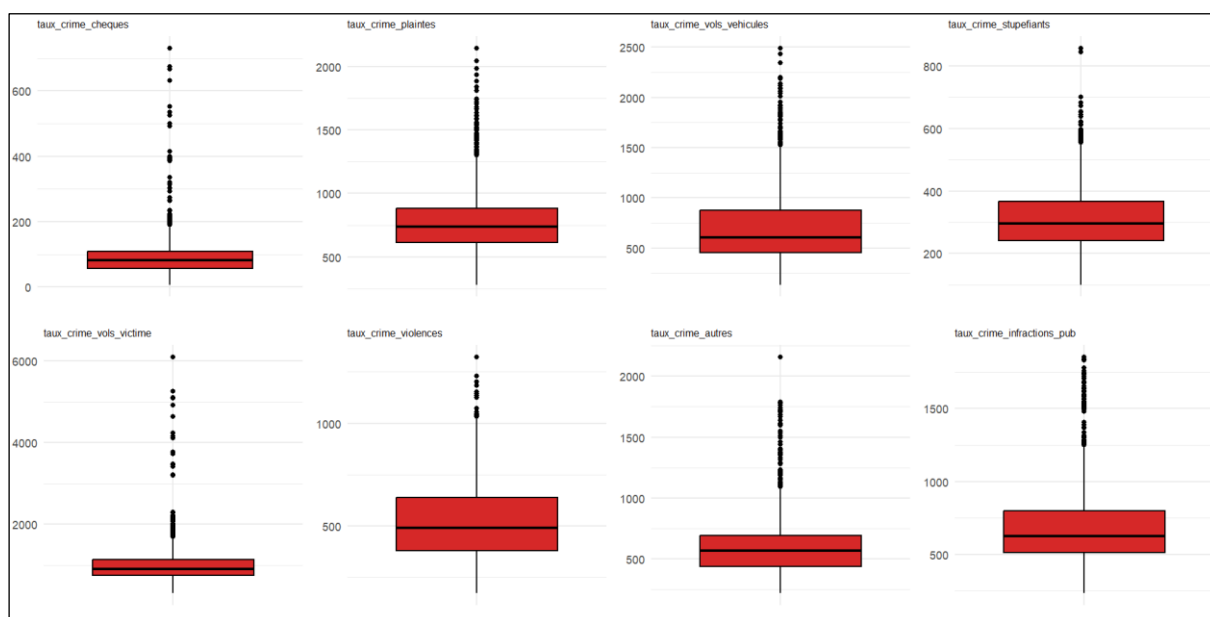
### Annexe 4 : Histogramme et boîte à moustache de la distribution du taux d'immigration



## Annexe 5 : Histogrammes des indicateurs de criminalité



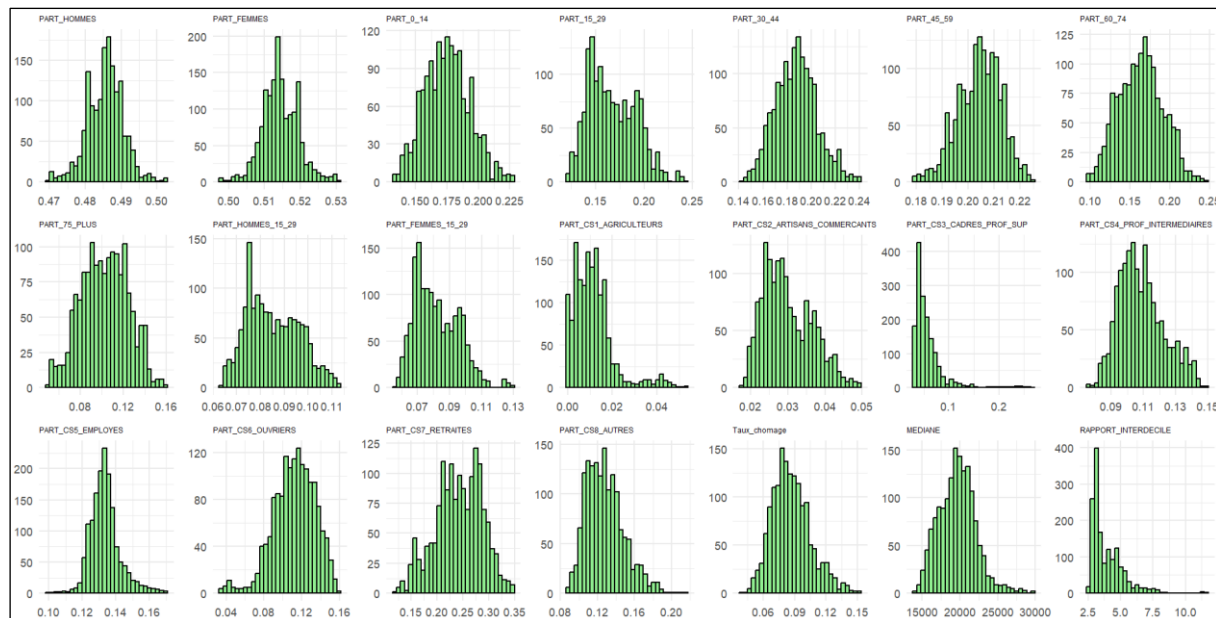
## Annexe 6 : Boîtes à moustaches des indicateurs de criminalité



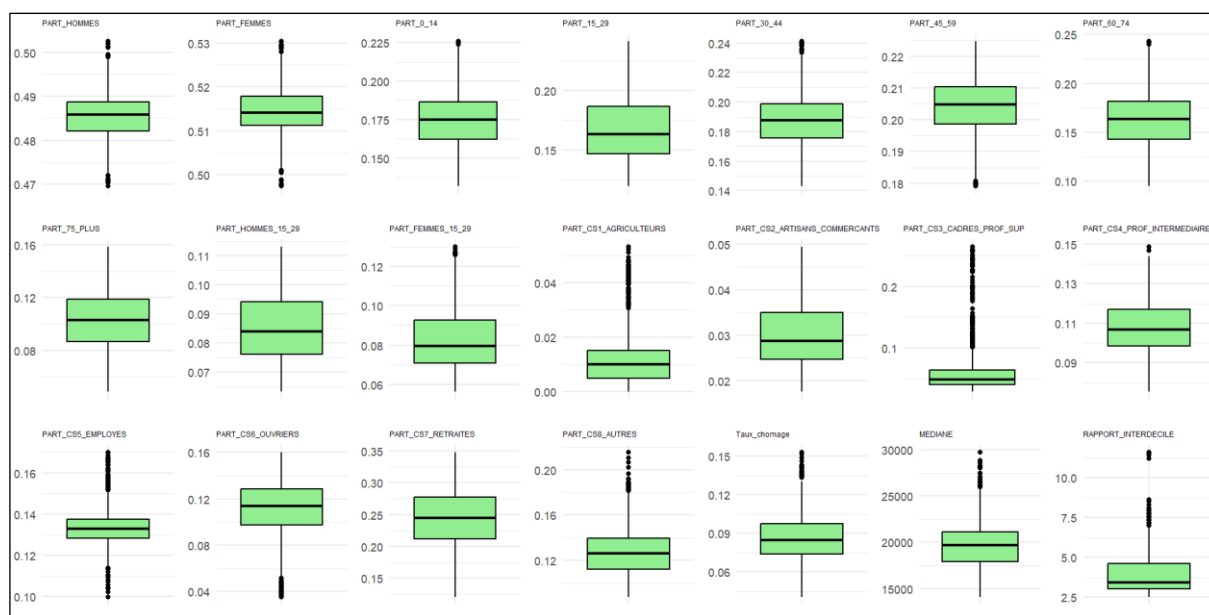
## Annexe 7 : Résumé des caractéristiques des distributions des indicateurs de criminalité et du taux d'immigration

TAUX_IMMIGRATION	taux_crime_cheques	taux_crime_plaintes	taux_crime_vols_vehicules
Min. : 0.01562	Min. : 7.137	Min. : 280.8	Min. : 135.7
1st Qu.: 0.04457	1st Qu.: 57.342	1st Qu.: 610.6	1st Qu.: 456.4
Median : 0.06436	Median : 82.119	Median : 738.6	Median : 604.5
Mean : 0.07389	Mean : 90.855	Mean : 772.0	Mean : 707.3
3rd Qu.: 0.09126	3rd Qu.: 109.233	3rd Qu.: 883.6	3rd Qu.: 878.8
Max. : 0.31143	Max. : 730.514	Max. : 2144.8	Max. : 2489.8
taux_crime_stupefiants	taux_crime_vols_victime	taux_crime_violences	
Min. : 100.9	Min. : 317.8	Min. : 174.1	
1st Qu.: 242.1	1st Qu.: 758.1	1st Qu.: 380.9	
Median : 296.5	Median : 901.4	Median : 492.2	
Mean : 312.0	Mean : 1008.1	Mean : 521.4	
3rd Qu.: 366.8	3rd Qu.: 1133.2	3rd Qu.: 638.5	
Max. : 855.9	Max. : 6103.6	Max. : 1323.6	
taux_crime_autres	taux_crime_infractions_pub		
Min. : 221.1	Min. : 233.9		
1st Qu.: 440.8	1st Qu.: 511.8		
Median : 567.6	Median : 626.2		
Mean : 605.5	Mean : 684.2		
3rd Qu.: 694.6	3rd Qu.: 800.2		
Max. : 2158.9	Max. : 1855.4		

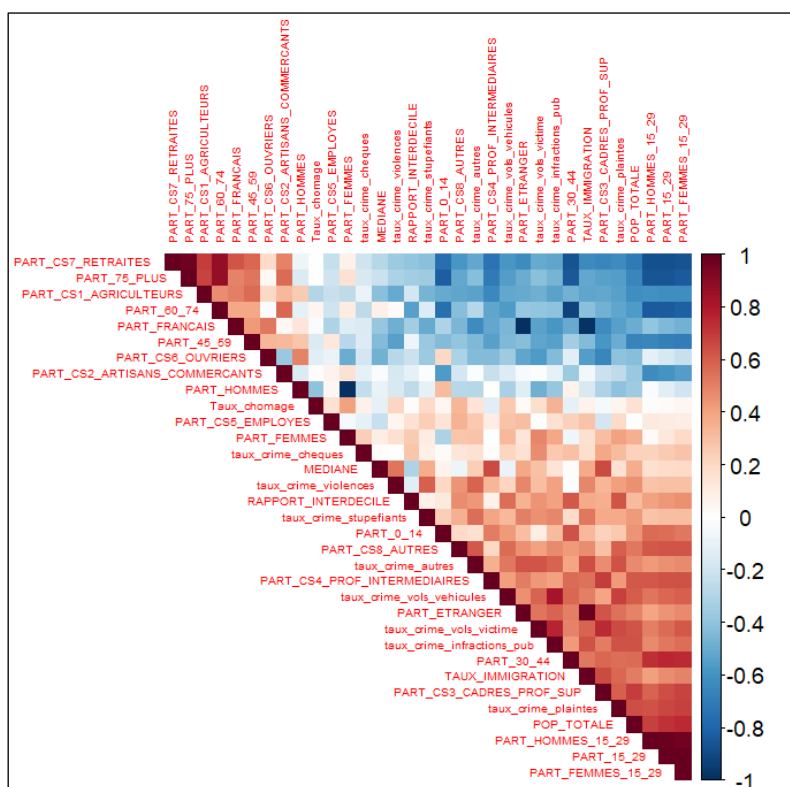
## Annexe 8 : Histogrammes des variables de contrôles



## Annexe 9 : Boîtes à moustache des variables de contrôles

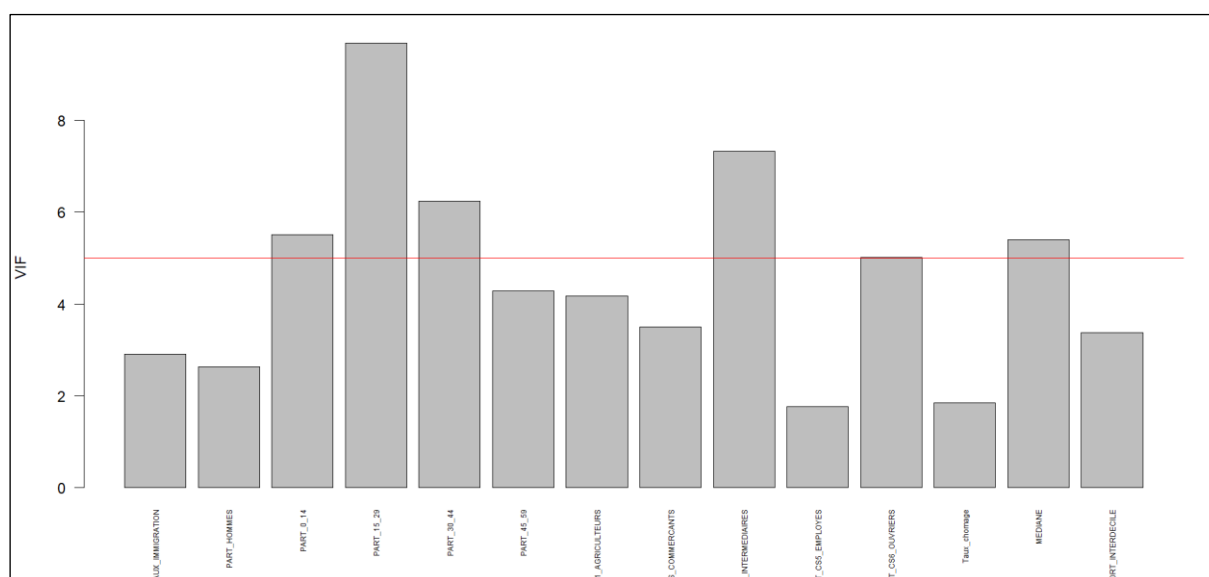


## Annexe 10 : Matrice de corrélation





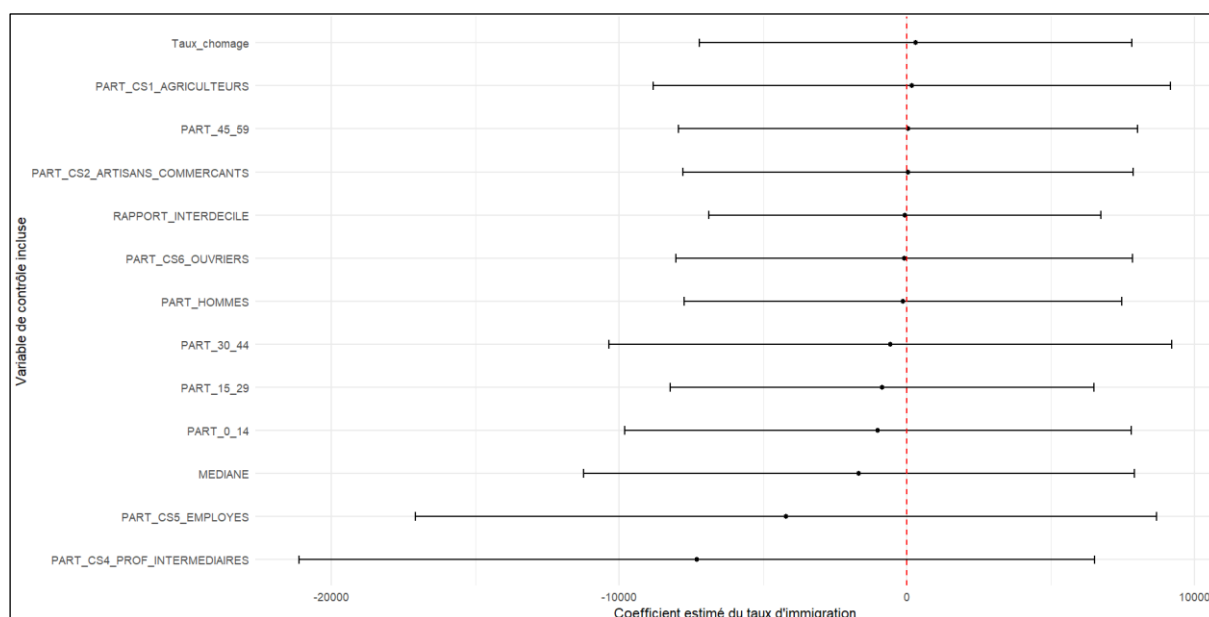
### Annexe 11 : Graphique du test de multicollinéarité (VIF)



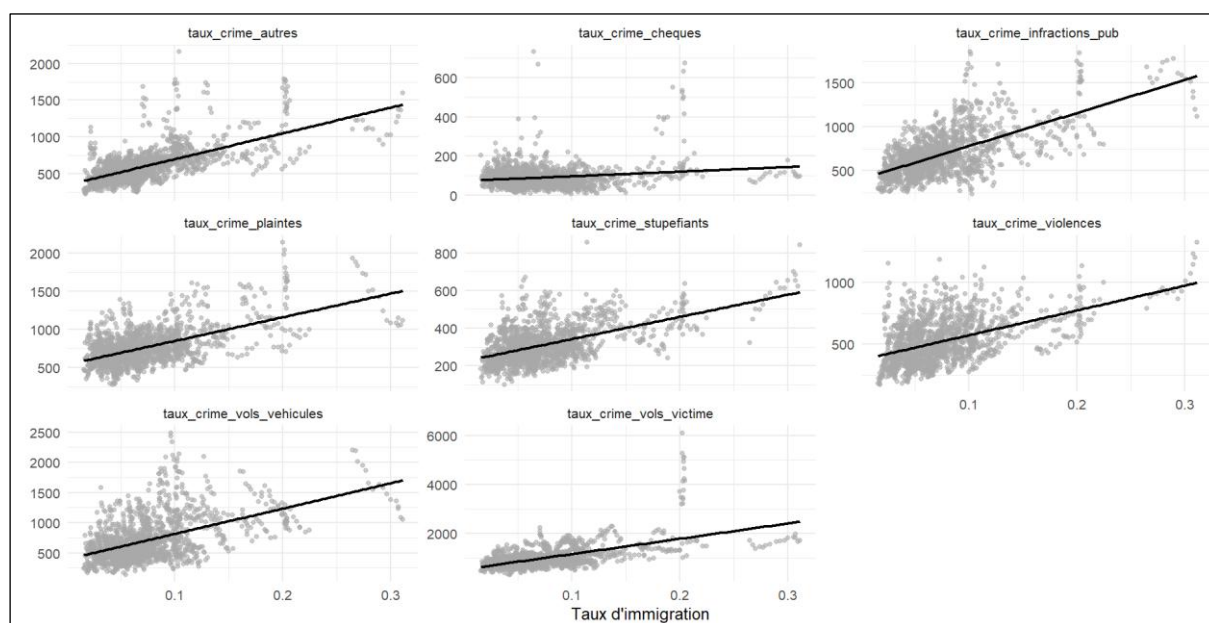
### Annexe 12 : Résultats des Test d'homoscédasticité de Breusch-Pagan estimé sur des modèles en effets fixes simples

Table: Résultats du test de Breusch-Pagan pour chaque régression				
Variable_dependante	Statistique	Df	P_value	Heteroscedasticite
:-----	:-----	:--	:-----	:-----
taux_crime_cheques	40.311	1	0.0000	Oui
taux_crime_plaintes	132.413	1	0.0000	Oui
taux_crime_vols_vehicules	40.281	1	0.0000	Oui
taux_crime_stupefiants	4.103	1	0.0428	Oui
taux_crime_vols_victime	140.880	1	0.0000	Oui
taux_crime_violences	1.109	1	0.2922	Non
taux_crime_autres	55.913	1	0.0000	Oui
taux_crime_infractions_pub	111.021	1	0.0000	Oui

### Annexe 13 : Exemple de l'effet estimé du taux d'immigration avec l'inclusion progressive des variables de contrôle sur le taux de vols sur victimes



### Annexe 14 : Relation linéaire simple entre le taux d'immigration et les différents types de criminalité



Annexe 15 : Résumer des régressions simples pour les différents taux de criminalités

	<i>Chèques</i>	<i>Plaintes</i>	<i>Vols sur véhicules</i>	<i>Stupéfiants</i>	<i>Vols (victime)</i>	<i>Violences</i>	<i>Autres délits</i>	<i>Infractions publiques</i>
(Intercept)	73.481***	542.610***	397.682***	224.896***	547.809***	373.389***	345.534***	405.330***
	(2.952)	(9.809)	(15.540)	(4.153)	(19.114)	(7.954)	(9.269)	(10.037)
<i>Taux d'immigration</i>	235.123***	3104.013***	4190.822***	1178.182***	6228.899***	2003.065***	3517.858***	3774.528***
	(34.286)	(113.932)	(180.491)	(48.240)	(222.000)	(92.386)	(107.651)	(116.576)
Num.Obs.	1536	1536	1536	1536	1536	1536	1536	1536
R2	0.030	0.326	0.260	0.280	0.339	0.235	0.410	0.406
F	47.027	742.260	539.123	596.499	787.257	470.083	1067.867	1048.344

. p < 0.1, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Annexe 16 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime cheques*

```
===== Résultats pour : taux_crime_cheques =====
OLS estimation, Dep. Var.: taux_crime_cheques
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate Std. Error   t value Pr(>|t|)
PART_HOMMES      -969.522582 2031.380819  -0.477273 0.640047
PART_15_29       3330.585039 1578.370367   2.110142 0.052052 .
PART_30_44      -1546.978155  941.917454  -1.642371 0.121306
PART_45_59       1815.836173 1256.783896   1.444828 0.169072
PART_CS4_PROF_INTERMEDIAIRES -5345.231054 2157.172745  -2.477887 0.025598 *
PART_CS5_EMPLOYES -5084.433343 1904.221014  -2.670086 0.017475 *
PART_CS6_OUVRIERS  -868.944510  996.397086  -0.872087 0.396906
MEDIANE           0.031616   0.018934   1.669857 0.115678
RAPPORT_INTERDECILE -19.323092 10.143093  -1.905049 0.076128 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 40.8      Adj. R2: 0.501802
                  Within R2: 0.158624
```

Annexe 17 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime vols vehicules*

```
===== Résultats pour : taux_crime_vols_vehicules =====
OLS estimation, Dep. Var.: taux_crime_vols_vehicules
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate Std. Error   t value Pr(>|t|)
TAUX_IMMIGRATION   -79.760246 3080.227067  -0.025894 0.9796830
PART_HOMMES        10256.104538 6506.181071   1.576363 0.1357944
PART_0_14          -1947.591059 3693.887498  -0.527247 0.6057331
PART_30_44         -6784.992556 4676.296636  -1.450933 0.1673894
PART_45_59         -7942.383437 3569.003383  -2.225379 0.0418168 *
PART_CS1_AGRICULTEURS -21170.646266 6195.947817  -3.416854 0.0038233 **
PART_CS2_ARTISANS_COMMERCANTS -8559.835816 8058.007316  -1.062277 0.3049141
PART_CS4_PROF_INTERMEDIAIRES -1639.865373 5791.309947  -0.283160 0.7809249
PART_CS6_OUVRIERS  -6269.035693 3534.612237  -1.773613 0.0964247 .
Taux_chomage       -4028.557321 1985.712773  -2.028771 0.0606195 .
MEDIANE             0.081594   0.032658   2.498402 0.0245832 *
RAPPORT_INTERDECILE  72.221133  30.772563   2.346933 0.0330748 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 93.5      Adj. R2: 0.928047
                  Within R2: 0.336535
```

Annexe 18 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime stupefiants*

```

===== Résultats pour : taux_crime_stupefiants =====
OLS estimation, Dep. Var.: taux_crime_stupefiants
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

```

	Estimate	Std. Error	t value	Pr(> t )
TAUX_IMMIGRATION	1331.257041	1238.444997	1.074942	0.299388
PART_HOMMES	3680.738094	4015.249407	0.916690	0.373809
PART_0_14	-1071.537551	1611.933349	-0.664753	0.516299
PART_15_29	544.477423	1426.893626	0.381582	0.708122
PART_30_44	-1939.746067	1606.035160	-1.207786	0.245829
PART_CS1_AGRICULTEURS	5301.773010	2726.336053	1.944651	0.070813
PART_CS2_ARTISANS_COMMERCANTS	-4645.920214	3136.777237	-1.481113	0.159276
PART_CS4_PROF_INTERMEDIAIRES	-90.130173	1856.851837	-0.048539	0.961927
PART_CS5_EMPLOYES	509.401697	1361.070156	0.374266	0.713445
PART_CS6_OUVRIERS	2233.952658	1338.266940	1.669288	0.115792
Taux_chomage	338.067170	510.374262	0.662391	0.517770
MEDIANE	0.007525	0.015687	0.479683	0.638372
RAPPORT_INTERDECILE	-1.393093	7.568388	-0.184067	0.856426

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 43.8      Adj. R2: 0.784718
                  Within R2: 0.036298

```

Annexe 19 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime vols victimes*

```

===== Résultats pour : taux_crime_vols_victime =====
OLS estimation, Dep. Var.: taux_crime_vols_victime
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

```

	Estimate	Std. Error	t value	Pr(> t )
TAUX_IMMIGRATION	-11747.122927	7005.155018	-1.676925	0.114267
PART_HOMMES	12491.171871	7497.316646	1.666086	0.116436
PART_0_14	2945.476171	3411.330095	0.863439	0.401492
PART_15_29	18235.620975	6573.543309	2.774093	0.014181 *
PART_30_44	3535.353262	3972.038409	0.890060	0.387486
PART_45_59	3935.019111	4321.820076	0.910500	0.376958
PART_CS1_AGRICULTEURS	-11334.040169	8464.867537	-1.338951	0.200526
PART_CS2_ARTISANS_COMMERCANTS	-1225.514089	8676.361269	-0.141247	0.889554
PART_CS4_PROF_INTERMEDIAIRES	-19694.855605	7970.416519	-2.470995	0.025948 *
PART_CS5_EMPLOYES	-10985.466892	6824.964634	-1.609601	0.128324
PART_CS6_OUVRIERS	-3434.061299	3386.839113	-1.013943	0.326687
Taux_chomage	-4634.398137	2752.043567	-1.683984	0.112874
MEDIANE	-0.004592	0.023671	-0.193991	0.848786

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 106.1     Adj. R2: 0.945232
                  Within R2: 0.222787

```

Annexe 20 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime violences*

```

===== Résultats pour : taux_crime_violences =====
OLS estimation, Dep. Var.: taux_crime_violences
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate   Std. Error   t value   Pr(>|t|)
TAUX_IMMIGRATION      -1618.668897   995.053450  -1.626716   0.12462
PART_HOMMES            3114.238664   2628.375400   1.184853   0.25451
PART_0_14              -174.506226   1483.417635  -0.117638   0.90792
PART_15_29             -962.895820   994.388213  -0.968330   0.34823
PART_30_44             -180.452032   1479.095420  -0.122002   0.90452
PART_45_59             -389.783235   1161.179424  -0.335679   0.74176
PART_CS1_AGRICULTEURS   2050.996772   2590.750797   0.791661   0.44090
PART_CS2_ARTISANS_COMMERCANTS -1344.153419   3017.859376  -0.445400   0.66239
PART_CS4_PROF_INTERMEDIAIRES -1643.311937   1208.261605  -1.360063   0.19390
PART_CS5_EMPLOYES       573.899173   1330.775362   0.431252   0.67242
Taux_chomage           739.662857   518.401621   1.426814   0.17412
MEDIANE                 0.006275    0.012178    0.515235   0.61390
RAPPORT_INTERDECILE     -2.969409    6.980940  -0.425359   0.67661
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 37.8      Adj. R2: 0.953441
                  Within R2: 0.030611

```

Annexe 21 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le *taux crime autres*

```

===== Résultats pour : taux_crime_autres =====
OLS estimation, Dep. Var.: taux_crime_autres
Observations: 1,536
Fixed-effects: CODDEP: 96, ANNEES: 16
Standard-errors: Clustered (CODDEP & ANNEES)

              Estimate   Std. Error   t value   Pr(>|t|)
TAUX_IMMIGRATION      2552.579145   2837.128029   0.899705   0.3824940
PART_HOMMES            6305.614748   6803.453094   0.926826   0.3686903
PART_15_29             -3551.488155   2174.817902  -1.633005   0.1232776
PART_30_44             -7867.695947   2283.121437  -3.446026   0.0036011 **
PART_45_59             -14790.390450   4180.624559  -3.537842   0.0029827 **
PART_CS1_AGRICULTEURS  -7402.697604   5588.173720  -1.324708   0.2050965
PART_CS2_ARTISANS_COMMERCANTS 2677.135415   9204.563373   0.290849   0.7751499
PART_CS4_PROF_INTERMEDIAIRES -9300.211775   3871.696491  -2.402102   0.0297026 *
PART_CS5_EMPLOYES      3107.649367   4382.378316   0.709124   0.4891237
MEDIANE                 0.041666    0.033734    1.235147   0.2357733
RAPPORT_INTERDECILE     60.796268    21.220565    2.864969   0.0118034 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 80.4      Adj. R2: 0.880836
                  Within R2: 0.247429

```

## Annexe 22 : Résumé des estimations pour les différentes régressions de la base initiale

Effet estimé du taux d'immigration selon le type de criminalité et le type de régression

Type_crime	LM simple	LM contrôle	FE (plm)	FEOLS	FE cluster plm	FEOLS cluster	FE stepwise	FE stepwise cluster
Chèques	235.123***	29.013	915.539*	915.539	915.539*	915.539	NA	NA
Plaintes	3104.013***	491.258***	-13392.791***	-13392.791***	-13392.791***	-13392.791***	-6983.302**	<b>-6983.302**</b>
Vols véhicules	4190.822***	679.122***	-8636.222***	-8636.222**	-8636.222***	-8636.222**	-79.76	<b>-79.76</b>
Stupéfiants	1178.182***	605.068***	343.488	343.488	343.488	343.488	1331.257	<b>1331.257</b>
Vols victime	6228.899***	2359.348***	27.349	27.349	27.349	27.349	-11747.123.	<b>-11747.123</b>
Violences	2003.065***	1574.736***	-1415.06***	-1415.06.	-1415.06***	-1415.06.	-1618.669	<b>-1618.669</b>
Autres	3517.858***	1628.259***	-1371.861.	-1371.861	-1371.861.	-1371.861	2552.579	<b>2552.579</b>
Infractions pub	3774.528***	2643.543***	-5004.14***	-5004.14***	-5004.14***	-5004.14**	-4043.738*	<b>-4043.738.</b>

## Annexe 23 : Résumé des estimations pour les différentes régressions de la base *Saphir*

Effet estimé du taux d'immigration selon le type de criminalité et le type de régression

Type_crime	LM simple	LM contrôle	FE (plm)	FE cluster plm	FE stepwise	FE stepwise cluster
Chèques	132.717	-62.994	-3046.431***	-3046.431***	NA	NA
Plaintes	3314.017***	2251.969***	-4715.409***	-4715.409***	-4112.953*	<b>-4112.953*</b>
Vols véhicules	4034.394***	3162.676***	-28188.614***	-28188.614***	3884.276	<b>3884.276</b>
Stupéfiants	4034.394***	996.642***	5727.505***	5727.505***	1356.625	<b>1356.625</b>
Vols victime	6343.513***	1928.356***	-1496.118.	-1496.118.	-1105.314	<b>-1105.314</b>
Violences	2496.253***	2382.696***	15436.401***	15436.401***	-466.594	<b>-466.594</b>
Autres	3715.107***	3077.294***	8143.318***	8143.318***	3298.805*	<b>3298.805.</b>
Infractions pub	3573.479***	2526.446***	-5242.617***	-5242.617***	109.919	<b>109.919</b>

## Annexe 24 : Résumé des estimations pour les différentes régressions de la base *Criminalité*

Effet estimé du taux d'immigration selon le type de criminalité et le type de régression

Type_crime	LM simple	LM contrôle	FE (plm)	FEOLS	FE cluster plm	FEOLS cluster	FE stepwise	FE stepwise cluster
Crimes violent	1725.75***	-723.396***	-1301.636	-1301.636	-1301.636	-1301.636	902.347	<b>902.347</b>
Escroqueries fraudes	1282.42***	382.057**	-3641.097***	-3641.097**	-3641.097***	-3641.097**	-3986.817**	<b>-3986.817**</b>
Vols dégradations	2932.352***	-1217.57***	-9286.271***	-9286.271**	-9286.271***	-9286.271*	-5018.536.	<b>-5018.536.</b>
Vols sans violence	6858.032***	-1849.256**	7746.041*	7746.041**	7746.041*	7746.041*	16913.474**	<b>16913.474*</b>
Vols véhicules	3346.259***	-682.759.	-5355.379*	-5355.379.	-5355.379*	-5355.379	408.548	<b>408.548</b>
Stupéfiants	1733.525***	-383.255	6415.319*	6415.319.	6415.319*	6415.319	336.173	<b>336.173</b>
Cambriolages	2652.127***	1208.309***	-5899.414*	-5899.414*	-5899.414*	-5899.414.	-1194.456	<b>-1194.456</b>



## Annexe 25 : Construction des différentes bases de données

Dans cette annexe sont présentés les différents traitements et constructions des bases de données utilisées dans les analyses. Pour ce faire, des scripts en Python ont été développés, des jointures et calculs ont été réalisés en langage R et certains traitements manuels ont été effectués sur Excel. L'ensemble des manipulations de données et des agrégations statistiques ont fait l'objet de vérifications systématiques, s'inscrivant dans une démarche rigoureuse visant à garantir la fiabilité et la cohérence des résultats obtenus. Afin de pouvoir reproduire les analyses réalisées, les liens des différentes bases de données utilisées sont disponibles dans la partie « *liens des bases de données* » dans la bibliographie.

### - Base de données initiale (2006-2021)

Afin d'exploiter les données des indicateurs de criminalité au niveau départementales, un traitement préliminaire a été effectué à l'aide du langage Python. Le fichier source utilisé, intitulé *tableaux-4001-ts.xlsx* et disponible sur le site de [data.gouv.fr](http://data.gouv.fr), contient de nombreuses feuilles Excel correspondant chacune à un département, ainsi que des feuilles agrégées pour la France entière et la France métropolitaine, qui ont été volontairement exclues de l'analyse.

Dans un premier temps, les bibliothèques nécessaires (à savoir *panda* et *re*) et le fichier Excel ont été importés. Un dictionnaire a ensuite été constitué pour faire correspondre les 107 index présents dans chaque feuille à leurs libellés exacts et à l'unité de compte associée (par exemple : « Victime », « Infraction », « Procédure », etc.). Pour faciliter l'analyse, les 107 types d'infractions recensés ont été regroupés en huit grandes catégories selon leur unité de compte, via une fonction de regroupement personnalisée. Les catégories résultantes sont les suivantes : « Chèques », « Plaintes », « Vols sur véhicules », « Stupéfiants et infraction par auteur », « Vols (Victime entendue) », « Violences physiques & psychologiques », « Autres délits & crimes », et « Violations institutionnelles et publiques ». Ce regroupement permet de simplifier l'analyse en réduisant la granularité des données tout en conservant la diversité des formes de criminalité.

Par la suite, le script a parcouru l'ensemble des feuilles correspondant aux départements. Pour chacune, les données ont été nettoyées (suppression des lignes et colonnes vides) et



seules celles contenant une colonne « Index » ont été retenues, afin de garantir l'uniformité des structures de données traitées. Le code détecte ensuite les colonnes correspondant aux années, en repérant celles dont le nom est un entier à quatre chiffres (par exemple : « 2016 », « 2017 », etc.). Ces colonnes sont utilisées pour transformer la base au format long à l'aide de la fonction « *melt* », ce qui permet de faire apparaître chaque observation comme une combinaison unique d'indicateur, d'année et de valeur. Le résultat de chaque feuille est stocké dans une liste, permettant ensuite de concaténer l'ensemble des départements. Enfin, si une erreur survient lors du traitement d'une feuille, celle-ci est signalée sans interrompre le traitement global.

Une fois toutes les feuilles valides traitées, le script concatène l'ensemble des données départementales à l'aide de la fonction « *pd.concat* », formant ainsi une base complète regroupant les huit indicateurs enregistrés par département et par année. Le code supprime également les doublons éventuels afin d'éviter les redondances susceptibles d'affecter les résultats économétriques. Enfin, la base consolidée est exportée au format Excel, permettant ainsi son exploitation dans le logiciel de traitement statistique Rstudio.

Concernant la base de données des estimations du recensement de la population (IRIS), cette base a également reçu un traitement par le logiciel Python. Avant toute chose, afin de construire cette base de données, l'ensemble des fichiers Excel (correspondant à une année d'estimation du recensement IRIS) ont été téléchargé et récupéré. Ensuite, pour chaque année, le fichier Excel correspondant est automatiquement localisé dans un dossier défini, en tenant compte du format de fichier .xls jusqu'en 2016, puis .xlsx à partir de 2017. Une fois chaque fichier chargé, les noms de colonnes dynamiques sont standardisés afin de garantir une homogénéité sur l'ensemble des années. Certaines colonnes descriptives non nécessaires à l'analyse (telles que les noms de commune, types IRIS, libellés, etc.) sont supprimées pour ne conserver que les variables quantitatives. Après avoir vérifié la présence de la variable départementale DEP, les données sont agrégées par département, permettant ainsi de passer de l'échelle IRIS à l'échelle départementale. Cette agrégation est répétée pour chaque année, et les résultats sont concaténés dans un seul tableau final. Une vérification est effectuée à chaque étape pour s'assurer que les agrégations produisent bien des résultats valides. Enfin, la base finale complète est sauvegardée au format Excel, rendant les données exploitables pour des analyses économétriques sur le logiciel Rstudio.

- Base de données *Saphir* (1999, 2010, 2015, 2021)

La base de données *Saphir*, issue de l'INSEE, constitue un recensement individuel massif avec sur quatre années de données et avec environ 50 millions d'observations initiales, chaque ligne représentant un individu. L'ensemble du traitement de cette base a été réalisé sous le logiciel R. Afin d'alléger la base et de ne conserver que les variables strictement nécessaires à l'analyse, plusieurs colonnes relatives à la géographie régionale (départements ou régions de résidence, de travail ou de naissance) ainsi que le statut conjugal ont été supprimées. Un travail de typage a été effectué pour convertir les variables dans un format adapté à l'analyse (facteurs, entiers, numériques). La variable diplôme a notamment été harmonisée pour regrouper les niveaux de qualification selon cinq grandes catégories allant de l'absence de diplôme jusqu'au supérieur. Après avoir exclu les individus non exploitables (élèves, mineurs, etc.), les variables sociodémographiques principales ont été agrégées à l'échelle départementale et annuelle à l'aide de la pondération fournie par l'INSEE, permettant de calculer la structure de population selon le sexe, la CSP, le niveau de diplôme, la nationalité, l'activité et les tranches d'âges. Des traitements complémentaires ont permis de supprimer les catégories obsolètes (notamment les militaires du contingent ou les secteurs d'activités « NES4 » redondants avec la CSP) et de neutraliser les valeurs manquantes restantes. La création du taux d'immigration a été réalisé selon la définition d'un immigré par l'INSEE. Enfin, les tranches d'âges ainsi que l'ensemble des variables ont été exprimées en proportions pour homogénéiser toutes les variables explicatives sous forme de parts démographiques, assurant ainsi la cohérence des unités dans la base harmonisée finale. Pour finir, diverses jointures (notamment avec les indicateurs de criminalité, le taux de chômage ou encore les indicateurs de richesses) entre les différentes bases de données ont été réalisées afin de créer une unique base exploitable dans les analyses.

- Base de données *Criminalité* revisité (2016-2021)

Cette base issue de *data.gouv.fr* et traitée par le SSMSI, a uniquement reçu un traitement manuel sur Excel afin d'organiser les différents indicateurs de criminalité en données de panel (une ligne égale un département et une année). Ensuite, les 18

indicateurs initialement présents ont été regroupés en fonction de leur unité de compte, c'est-à-dire selon le type de victimes concernées ou la localisation des faits. Cette opération d'agrégation a permis de les harmoniser en sept indicateurs de criminalité, correspondant à des groupes hétérogènes mais cohérents sur le plan statistique. Enfin, sur R, les taux pour 100 000 habitants et pour 100 000 logements (pour le taux de cambriolages) ont été calculés.

## Table des matières

Glossaire sigles .....	2
Introduction .....	3
I - Analyse et revue de la littérature .....	6
II - Présentation des variables d'intérêt.....	9
A) Taux d'immigration .....	9
B) Taux de criminalité .....	12
II - Présentation des variables de contrôles.....	16
a) Variables démographiques .....	16
b) Variable socio-professionnelle .....	17
c) Variables économiques.....	17
III - Modèles économétriques utilisées.....	19
IV - Méthodologie appliquée .....	24
a) Méthodes d'estimations et bases de données.....	24
b) Analyse descriptive des variables des modèles finaux.....	27
c) Tests de validation économétriques.....	29
V - Analyse des résultats.....	31
VI - Analyse des estimations des bases alternatives .....	37
1) Base de données du recensement harmonisé ( <i>Saphir</i> ) .....	37
2) Base de données des indicateurs de criminalité revisité depuis 2016 .....	40
VII - Résumés des résultats finaux.....	45
VIII - Discussion.....	46
a) Analyse des résultats .....	46
b) Limites .....	48
c) Recommandations .....	49
Conclusion .....	50
Bibliographie, Sitographie, liens des bases de données.....	52
Bibliographie .....	52
Sitographie.....	53
Liens des bases de données .....	54
Base initiale (2006-2021) .....	54
Base de données <i>Saphir</i> (1999, 2010, 2015, 2021) .....	54
Base de données <i>Criminalité</i> revisité (2016-2021) .....	54
Variables de contrôles.....	55
Packages Python et R.....	55

<b>Cours Master ECAP .....</b>	<b>55</b>
<b>ANNEXES.....</b>	<b>56</b>
Annexe 1 : Évolution du taux d'immigration et du nombre d'immigrés de 2006 à 2021 en France.....	56
Annexe 2 : Évolution des indicateurs du taux de criminalité par années de 2006 à 2021 ....	56
Annexe 3 : Évolution du taux de chômage moyen en France Métropolitaine de 1982 à 2024 .....	57
Annexe 4 : Histogramme et boîte à moustache de la distribution du taux d'immigration ....	57
Annexe 5 : Histogrammes des indicateurs de criminalité .....	58
Annexe 6 : Boîtes à moustaches des indicateurs de criminalité .....	58
Annexe 7 : Résumé des caractéristiques des distributions des indicateurs de criminalité et du taux d'immigration .....	59
Annexe 8 : Histogrammes des variables de contrôles.....	59
Annexe 9 : Boîtes à moustache des variables de contrôles.....	60
Annexe 10 : Matrice de corrélation .....	60
Annexe 11 : Graphique du test de multicolinéarité (VIF).....	61
Annexe 13 : Exemple de l'effet estimé du taux d'immigration avec l'inclusion progressive des variables de contrôle sur le taux de vols sur victimes .....	62
Annexe 14 : Relation linéaire simple entre le taux d'immigration et les différents types de criminalité .....	62
Annexe 15 : Résumer des régressions simples pour les différents taux de criminalités.....	63
Annexe 16 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime cheques</i> .....	64
Annexe 17 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime vols vehicules</i> .....	64
Annexe 18 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime stupefiants</i> .....	65
Annexe 19 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime vols victimes</i> .....	65
Annexe 20 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime violences</i> .....	66
Annexe 21 : Résultats de la régression finale en effets fixes avec variables de contrôles et clusters pour le <i>taux crime autres</i> .....	66
Annexe 22 : Résumé des estimations pour les différentes régressions de la base initiale ....	67
Annexe 23 : Résumé des estimations pour les différentes régressions de la base <i>Saphir</i> .....	67
Annexe 24 : Résumé des estimations pour les différentes régressions de la base <i>Criminalité</i> .....	67
<b>Annexe 25 : Construction des différentes bases de données .....</b>	<b>68</b>
- Base de données initiale (2006-2021).....	68

- Base de données Saphir (1999, 2010, 2015, 2021).....	70
- Base de données <i>Criminalité</i> revisitée (2016-2021).....	70
<b>Table des matières .....</b>	<b>72</b>