

**Proposition de communication pour les
2e Rencontres Francophones Transport Mobilité (RFTM)
Montréal, 11-13 juin 2019**

Titre :

Peut-on mesurer la fraude sans enquête ? Résultats d'une analyse comparative des données du réseau de transports publics de Lyon.

Auteur(s) :

Oscar EGU,
Keolis Lyon, 19 Boulevard Marius Vivier Merle, 69003 Lyon
oegu@keolis-lyon.fr

Patrick Bonnel,
LAET - ENTPE, Rue Maurice Audin, 69518 Vaulx-en-Velin Cedex,
Patrick.Bonnel@entpe.fr

Mots-clés :

Transport public, Évasion tarifaire, Collecte de données, big data, Lyon

Résumé :

Dans un contexte d'urbanisation croissante et de prise de conscience des enjeux écologiques, il est très probable que les transports en commun auront un rôle important à jouer dans les décennies à venir. L'amélioration de la qualité et de l'attractivité de l'offre ainsi que des financements adéquats sont indispensables. Au regard de ces éléments, l'évasion tarifaire (ou fraude) c'est-à-dire le fait d'utiliser un service de transport sans s'acquitter des droits adéquats ou sans respecter la réglementation peut être considérée comme un risque qu'il faut quantifier et maîtriser (Reddy et al., 2011 ; Barabino et al. 2015 ; Delbosc et Currie 2018). Malheureusement, ce phénomène est difficile à estimer (Reddy et al., 2011 ; Dauby et Kovacs, 2007 ; Lee 2011). C'est pourquoi l'évasion tarifaire est traditionnellement mesurée à partir d'enquêtes de terrain qui doivent fournir des estimations globales fiables, mais avec des niveaux de précisions spatiales et temporelles limitées. Les avancées récentes dans les systèmes de transports intelligents, et dans la collecte automatique de données, offrent de nouvelles perspectives pour l'analyse de l'évasion tarifaire (Delbosc et Currie 2018 ; Pourmonet et al. 2015). L'objet de cette communication est donc d'explorer les possibilités de mesure de la fraude à l'aide de ces nouvelles sources de données et notamment de répondre à la question :

- Peut-on mesurer la fraude de manière satisfaisante sans enquête ?

Après une rapide revue de la littérature, la première partie de la présentation vise à détailler le cas d'étude et les données disponibles. Deux modalités d'accès au réseau coexistent sur le réseau de Lyon. Le réseau de métro peut être considéré comme un environnement fermé. Inversement, le réseau de tramways et de bus fonctionne comme un système ouvert (« open proof of payment » en anglais). L'ensemble du réseau est équipé d'un système billettique qui date de 2002, et qui permet à la fois l'utilisation de cartes à puce et de tickets magnétiques papiers. Pour le métro, l'accès aux stations se fait par des portiques équipés de barrières anti-

intrusion et d'un système de comptage bidirectionnel. Pour ouvrir les barrières et accéder aux quais, il est nécessaire de valider un titre de transport valable. Dans les tramways et les bus, des systèmes de comptage sont installés au niveau de chaque porte et des valideurs embarqués sont disponibles à l'intérieur des véhicules. La réglementation du réseau impose la validation systématique d'un titre de transport, à chaque montée dans un véhicule. Toutefois, cette contrainte n'est pas systématiquement respectée. L'opérateur en charge du réseau classe donc les irrégularités tarifaires en deux grandes familles :

- Irrégularité tarifaire avec perte de recette ;
- Irrégularité tarifaire sans perte de recette.

Cette classification est essentielle, car elle implique des irrégularités de comportement différentes, et il est important pour l'opérateur de pouvoir les quantifier séparément. Les deux types d'irrégularités ont un impact important sur la représentativité des données billettiques, mais seul le premier type d'irrégularité a un impact sur les recettes de l'opérateur. Pour s'assurer du respect de la réglementation, en moyenne 15 000 contrôles sont réalisés par jour, par des contrôleurs assermentés qui peuvent aussi dresser des procès-verbaux. Les contrôleurs utilisent un appareil portable qui permet de lire les titres, vérifier leur validité et administrer des amendes. Les données récoltées par ces appareils de contrôle sont consolidées dans le système d'information de l'opérateur. Nous avons à notre disposition trois sources de données, collectées automatiquement, provenant du système de comptage, du système billettique et du système de contrôle des titres. En plus de ces données, chaque année, trois vagues d'enquêtes fraude sont réalisées par un prestataire indépendant pour un total d'environ 57 000 personnes enquêtées. Ces enquêtes sont par sondage aléatoire avec une stratification par mode type de jour et période. Elles sont considérées comme la source de données la plus fiable concernant le niveau de fraude.

La deuxième partie de la présentation fournit un aperçu du traitement des données et des méthodes pour calculer les différents niveaux d'irrégularités tarifaires à partir des différentes sources de données disponibles. Les méthodes consistent à calculer des ratios. Par exemple, un niveau total d'irrégularités peut être estimé en divisant, pour une période et un espace donné, la quantité de validations par la quantité de personnes comptées. De même, les données issues des équipements de contrôle des titres permettent de quantifier à la fois le nombre de personnes inspectées, le type de titre inspecté et si une irrégularité est détectée, la nature de l'irrégularité. En combinant ces données avec les données des procès-verbaux, il est donc possible de déterminer, parmi les personnes contrôlées, la proportion totale d'irrégularités et la proportion d'irrégularités avec perte de recette. Les données de l'enquête permettent aussi de quantifier, avec des intervalles de confiance, à la fois la part totale d'irrégularités ainsi que la part d'irrégularités avec perte de recette.

La troisième partie de la présentation a pour objectif de détailler et discuter les résultats de cette étude comparative. Nous commençons par une présentation des résultats par mode de transport (bus, tramway, métro et funiculaire). Elle permet d'illustrer certaines différences importantes entre les sources de données, mais aussi selon les modes. L'analyse par ligne et par type de jour permettent de confirmer et d'affiner les premières observations. Plus précisément, l'analyse comparative permet de constater que les données récoltées par les contrôleurs sous-estiment systématiquement les niveaux d'irrégularités, que ce soit avec ou sans évasion de recette. L'analyse comparative par jour et par ligne montre que la structure de la fraude n'est pas bien représentée par ces données. De ce fait, nos résultats nous laissent penser que les données récoltées par les contrôleurs ne sont pas exploitables pour mesurer de manière efficace le niveau de fraude, et cela pour plusieurs raisons qui sont détaillées dans la présentation. L'utilisation du ratio entre le nombre de validations et le nombre de personnes

comptées est une piste d'analyse plus prometteuse. Dans le cas du métro, ce ratio se rapproche des enquêtes. Les différences marginales peuvent s'expliquer par le fait que le système de comptage ne permet pas de comptabiliser l'ensemble des comportements irréguliers, par exemple l'utilisation de tickets qui expirent une fois à l'intérieur du réseau de métro. Dans le cas, des systèmes ouverts comme le bus et le tramway, la comparaison des sources de données indique que le ratio validation/comptage surestime le taux d'irrégularités, mais les variations par ligne et par jour sont bien corrélées avec l'enquête et permettent d'envisager des clefs de passage ou l'exploration de la structure de la fraude. Les éléments permettant d'expliquer les différences entre les enquêtes et le ratio validation/comptage seront aussi discutés durant la présentation.

Les limites de cette recherche sont abordées en conclusion ainsi que les perspectives d'approfondissement. Nous insistons sur l'importance de la qualification des nouvelles sources de données à travers la comparaison avec des sources de données traditionnelles. À cet égard, les résultats de cette recherche sont intéressants. Ils permettent aux opérateurs, désireux d'utiliser ces nouvelles sources de données, de mieux saisir leur potentialité, mais aussi de prévenir des usages inappropriés.

Références :

Barabino, Benedetto, Sara Salis, and Bruno Useli. "What are the determinants in making people free riders in proof-of-payment transit systems? Evidence from Italy." *Transportation Research Part A: Policy and Practice* 80 (2015): 184-196.

Delbosc, Alexa, and Graham Currie. "Why do people fare evade? A global shift in fare evasion research." *Transport Reviews* (2018): 1-16.

Reddy, Alla, Jacqueline Kuhls, and Alex Lu. "Measuring and controlling subway fare evasion: improving safety and security at New York City transit authority." *Transportation Research Record: Journal of the Transportation Research Board* 2216 (2011): 85-99.

Dauby, Laurent, and Zoltan Kovacs. "Fare evasion in light rail systems." *Transportation Research E-Circular E-C112* (2007).

Lee, Jason. "Uncovering San Francisco, California, Muni's proof-of-payment patterns to help reduce fare evasion." *Transportation Research Record: Journal of the Transportation Research Board* 2216 (2011): 75-84.

Pourmonet, Hugo, Samuel Bassetto, and Martin Trépanier. "Vers la maîtrise de l'évasion tarifaire dans un réseau de transport collectif." *11e Congrès International De Génie Industriel-CIGI2015* (2015).