

DIALOGUE

La lettre d'information de DIAL

N° 53

Avril 2019

Focus : TOFLIT18, le retour.

Présenté dans le Dialogue n°43 de juillet 2016, l'objectif du projet TOFLIT18 est d'ouvrir de nouvelles voies de recherche sur l'histoire économique française à la veille de la révolution industrielle grâce aux données des archives du Bureau de la Balance du Commerce (L. Charles et G. Daudin, 2011) portant sur plus de 500 000 flux du commerce extérieur français.

Trois ans plus tard, nous exposons ici un bref aperçu de l'évolution du projet et de sa situation actuelle. Pour plus de détails, les lecteurs intéressés ou ceux entreprenant des projets similaires pourront consulter le site <https://toflit18.hypotheses.org>.

La principale vocation du projet était d'enrichir les connaissances sur les spécialisations commerciales locales et nationales françaises, tant sur leurs structures que leurs évolutions. Il avait également pour but d'analyser les effets de l'évolution des politiques économiques mercantilistes (notamment l'effet de *l'Exclusif* qui voulait réserver le commerce entre les colonies aux navires français) et des guerres sur le commerce et l'économie française afin de mieux appréhender l'interaction entre les événements politiques et le commerce international.

Une telle entreprise nécessitait d'associer des données nationales (macroéconomiques) et locales. Les données quantitatives produites par le Bureau de la Balance du Commerce (L. Charles et Daudin 2011) offrent cette opportunité grâce à la part importante de données macroéconomiques et celle, non moins significative, de données locales disponibles aux archives régionales françaises (ex : Bayonne, Bordeaux, La Rochelle, Marseille, Nantes, Rouen). L'analyse de ces données permet ainsi d'approfondir nos connaissances sur les phénomènes économiques de développement et de transformation économique mais aussi sur les déterminants du commerce international. Préalablement et afin de rendre ces données utilisables à des fins de recherche, un certain temps a été dédié à la collecte et au nettoyage des données existantes. Le premier traitement consistait à s'assurer que les données contenaient suffisamment d'informations fiables pour qu'elles soient exploitables en toute sécurité. Tout comme la fiabilité des données, il était nécessaire d'apprécier les difficultés relatives à l'adéquation des activités marchandes et des catégories administratives, à l'établissement d'une liste fiable et harmonisée des biens et des entités géographiques ainsi qu'au traitement des prix.



DIALOGUE
est une publication
de DIAL

Ce programme de recherche couvre une période cruciale : la révolution industrielle. Celle-ci établit les bases économiques de la France et de l'Europe et, surtout, déclenche leur entrée dans l'ère moderne. Une pièce essentielle du puzzle de l'histoire économique du XVIII^e siècle est le commerce international. Mais bien que le commerce extra-européen ait été relativement bien étudié, le commerce intra-européen reste, quant à lui, quelque peu négligé. Seules quelques études, partielles et fondées sur des méthodes économétriques hétérogènes, existent sur ce sujet. Cependant, leurs résultats sont souvent trop généraux et insuffisamment robustes pour permettre des analyses et des comparaisons économiques précises.

La rareté des travaux mobilisant les statistiques du commerce extérieur du début de la période moderne est regrettable, d'autant plus que les flux du commerce extérieur sont les données économiques que les premiers États modernes ont le mieux collectées. En effet, les premières tentatives de mesure du commerce extérieur remontent au début du XVII^e siècle. Des séries de données annuelles ont été localisées en Suède à partir de 1732, aux Pays-Bas autrichiens à partir de 1759, à Milan en 1762, 1769, 1778 et 1791, à Venise à partir de 1769-70, au Portugal en 1776-77, 1783 et 1789, et en Espagne à partir de 1778. La recherche sur le commerce des îles britanniques est, elle, bien plus avancée que dans les autres pays. Un ouvrage récent a renouvelé la recherche sur le commerce écossais au XVIII^e siècle (P. R. Rössner 2008) et, actuellement, une équipe coordonnée par David Jacks, Kevin O'Rourke et Alan Taylor recueille des données sur le commerce anglais et britannique de 1696 au début du XXI^e siècle.

Pour ce qui est du cas français, l'ampleur et la complexité des données collectées par l'État français mais aussi la dispersion d'une grande partie de leurs sources pendant la Révolution française, ont rendu leur traitement plus ardu. C'est pourquoi la plupart des études sur le commerce extérieur « national » français au XVIII^e siècle ont mobilisé une proportion limitée des archives. En effet, les études au niveau (des) pays n'ont jamais utilisé l'ensemble des données françaises collectées au niveau central. Certains travaux sur le commerce international, le transport maritime en France et au niveau local, en particulier dans les ports, comportent des réflexions sur l'effet économique du commerce extérieur mais ne fournissent pas de discussion approfondie à un niveau plus agrégé. De récents travaux de recherche ont néanmoins suscité un intérêt accru pour cette question et ont conduit à étudier de près certains échanges bilatéraux entre la France et d'autres pays/régions européennes. Cependant, ces travaux restent limités puisqu'ils ne développent pas de traitement quantitatif systématique de l'ensemble des échanges français et n'ont pas recours aux outils économétriques.

1. Comment le projet TOFLIT18 a-t-il été réalisé?

Le succès du projet est principalement imputable à la qualité scientifique et organisationnelle de l'équipe de recherche. Le travail s'est articulé autour de réunions mensuelles ou lors de conférences annuelles. La régularité des réunions a permis de discuter des questions cruciales, d'élaborer une méthodologie commune et d'approfondir la compréhension de chaque membre de l'équipe. De plus, chaque année, une confrontation des différents travaux à la communauté scientifique était proposée, permettant ainsi d'ouvrir des débats et de favoriser les discussions sur la méthodologie ou encore l'interprétation des premiers résultats.

La première étape du projet fut l'élaboration d'une base de données du commerce extérieur de la France grâce aux données du Bureau de la balance du commerce de 1716 à 1821 (intégrant des données locales de 1716 à 1781). Les données brutes étaient relativement cohérentes, ce qui a simplifié les discussions sur les questions conceptuelles. De plus, Guillaume Daudin et Loïc Charles, membres fondateurs de TOFLIT18, avaient déjà été confrontés à la collecte et à la retranscription de ces données. De même, l'équipe scientifique était constituée de membres impliqués dans d'autres projets de construction de bases de données. Un des principaux défis fut donc la quantité de travail nécessaire à la retranscription et au traitement des flux de données, ainsi qu'à la vérification de la fiabilité des données collectées.

Une autre difficulté portait sur la classification des marchandises. En effet, au vu de la couverture temporelle de cette base de données et des variations dans la dénomination des biens au fil du temps, un des principaux défis fut l'élaboration d'un tableau général de correspondance des appellations des biens à différentes périodes et lieux. Ainsi, pour répondre aux besoins scientifiques du projet, il fallut identifier et catégoriser les biens pour que les données puissent être harmonisées et, in fine, analysées. Il parût nécessaire de créer une interface d'utilisation commune pour que les données puissent s'adapter aux approches, potentiellement multiples, des chercheurs. Néanmoins, cette interface devait répondre à des critères de durabilité (c'est-à-dire qu'elle devait pouvoir être utilisée à l'avenir) et d'extension des thématiques de recherche (par exemple, pouvoir répondre à des questions de recherche sur le monde matériel du XVIIIe siècle jusqu'au début du XIXe siècle).

Afin de répondre à ces objectifs et grâce à l'expérience acquise lors du projet PSPN/Sciences Po Medialab, la méthodologie « *datascape* » fut appliquée. Cette méthodologie repose sur un panel de méthodes et d'outils des données numériques qui favorisent l'utilisation des analyses exploratoires (Tukey, 1977) dans la recherche en sciences sociales (Latour et al., 2011). Elle vise à fournir une méthode exploratoire d'analyses des données et modélise l'information provenant des archives à l'aide d'un outil de navigation. Pour ce faire, un processus ascendant traitant et classant les données en différentes catégories est utilisé. Plus précisément, les documents originaux sont numérisés (en prenant en compte les variations d'orthographe) puis, étape par étape, les données sont normalisées afin de diminuer et supprimer l'orthographe ou la diversité géographique, et ceci sans perte d'information. Cette méthode permet d'obtenir une solide base pour créer des classifications sur des questions de recherche spécifiques. Par exemple, afin d'étudier le marché international des médicaments, il est nécessaire d'identifier les noms des produits correspondant à des médicaments dans le contexte du XVIIIe siècle et de les dissocier des autres types de marchandises. A titre illustratif, le mercure était-il traité comme un produit médical ou non ? Afin de pouvoir répondre à cette question, il est essentiel de se renseigner sur ces différents thèmes et de mobiliser les savoirs des dictionnaires et encyclopédies du XVIIIe siècle.

2. Les principaux résultats de TOFLIT18

Le premier résultat de ce projet est donc la création d'une base de données unique sur le commerce international français du XVIIIe siècle. Celle-ci rassemble actuellement près de 515 000 flux commerciaux (vérifiés) auxquels s'ajouteront prochainement des dizaines de milliers d'autres. L'accès aux données est conservé dans un dépôt privé GitHub (https://github.com/medialab/toflit18_data) sous format csv et sera rendu public dès que la documentation sera achevée (fin 2019). Toute personne désirant y avoir accès est priée de s'adresser aux membres de TOFLIT18.

Une des réalisations les plus impressionnantes du projet est l'achèvement du *datascape* (<http://toflit18.medialab.sciences-po.fr/#/home>). Il s'agit d'un outil permettant d'explorer un grand nombre de données hétérogènes. Le glossaire des produits explore la signification de plus de 2 000 biens différents. La classification des biens permet de répertorier dans 25 000 catégories plus de 55 000 dénominations de biens sans perte d'information. De plus, en s'inspirant de la SITC, le nombre de catégories est réduit à 25, ce qui permet de faciliter le point d'entrée pour les chercheurs. Le *datascape* offre trois points de vue différents sur le commerce français (séries chronologiques, réseaux des produits et des sites), et deux vues supplémentaires pour les données françaises (métadonnées et classifications). Le code est disponible sous la licence en libre accès : <https://github.com/medialab/toflit18>.

Ces deux réalisations majeures ont nécessité un temps considérable, initialement alloué aux travaux d'analyses. Malgré cela, un premier ouvrage a été publié et d'autres sont à venir. Le numéro des statistiques comparatives sur le commerce international (Loïc Charles et Daudin, 2015) a été qualifié d'« ouvrage de référence pour quiconque s'intéresse au début du commerce moderne » (O'Rourke 2017). Celui-ci est disponible en accès libre sur <https://www.ofce.sciences-po.fr/publications/srevue.php?num=140>. Prochainement, un chapitre portant sur les conséquences

économiques des guerres françaises sera publié dans un ouvrage collectif dirigé par Patrick O'Brien (LSE). En parallèle, un certain nombre d'études comparatives ont été publiées ou sont actuellement au stade de document de travail. La comparaison de ces données françaises à des sources de données comparables comme celles de la navigation du détroit du Sound, et du commerce d'Hambourg et de la Belgique est à un stade avancé et démontre qu'il existe de solides fondations à la réalisation de travaux économiques et historiques, sous condition d'une utilisation prudente de ces données. En particulier, une version finale du papier sur le détroit du Sound a été écrite et devrait être publiée prochainement dans une revue anglaise.

Pour ce qui est de la version française, une première ébauche présentant le datascape et l'évolution de la classification des biens dans la statistique française a été présentée lors de séminaires et conférences. Prochainement, une version française de ce travail méthodologique sera publiée dans la Revue de Synthèse (début 2019). Des travaux sur le commerce des médicaments, l'effet de la perte du Canada, l'effet de la seconde guerre de cent ans ainsi que sur le fonctionnement du commerce du grain ont été récemment entrepris.

Accessoirement, deux réalisations imprévues ont également été menées. Tout d'abord, une vingtaine de stagiaires de tous les niveaux de qualification (L3 à M2) a été initiée à la recherche. Cette expérience a même suscité l'envie, chez certains, de poursuivre une carrière académique (dont un en histoire économique). Enfin, nous avons entièrement documenté l'avancement du projet sur un blog Hypothèses (<https://toflit18.hypotheses.org>). Ce travail supplémentaire de vulgarisation et de suivi permet l'accès aux procès-verbaux des réunions mensuelles de l'équipe, ce qui représente une source d'informations non-négligeable (notamment sur les difficultés à traverser) pour toutes personnes souhaitant entreprendre un tel projet.

A travers ses différents outils, TOFLIT18 constitue une véritable infrastructure pour la recherche et représente un tremplin vers la création d'un réseau européen de recherches similaires.

3. Quelles sont les principaux enseignements du projet TOFLIT18 ?

Ce projet a pu voir le jour grâce aux financements que nous avons réussi à obtenir (à hauteur de 72% du montant souhaité initialement). Bien que reconnaissants des fonds octroyés, le fait que la totalité du projet n'ait pas été financée a amené l'équipe à prendre un certain nombre de décisions imprévues qui ont ralenti l'avancement du projet. A titre d'exemple, la collecte de données devait initialement se faire au moyen de transcripteurs professionnels. Du fait d'un budget réduit, la mobilisation de ces derniers n'était pas envisageable et des stagiaires ont été employés. N'y ayant pas songé auparavant, des difficultés pour le recrutement comme pour leur encadrement ont été rencontrées (les chefs de projet ne pouvant encadrer plus de 4 stagiaires à la fois). Les résultats furent globalement de bonne qualité mais en raison d'une hétérogénéité des productions des stagiaires mobilisés, des vérifications approfondies s'avèrent nécessaires. De plus, la restriction budgétaire (de quasiment 100 000 euros) rendit impossible l'embauche pendant un an et demi d'un ingénieur d'étude permanent qualifié, paléographe (qui aurait pu vérifier les transcriptions) et techniquement compétent (nous aurions toutefois pu le former). Ce recrutement aurait permis d'avoir un individu en charge de gérer le flux de données des archives originales aux données du datascape à travers l'intégration d'images, des transcriptions originelles et corrigées.

Par conséquent, la majeure partie du travail de conservation a été effectuée par un des chefs d'équipe (professeur et chercheur à temps plein) aidé occasionnellement par d'anciens stagiaires embauchés sur de courtes périodes (1 à 6 mois). Inévitablement, ce travail n'a pas été achevé dans les temps voulus et est toujours en cours. Cet épisode imprévu a eu deux conséquences : le programme d'interprétation des données a été retardé (puisque'il faut pouvoir disposer de tout le panel de données avant d'entreprendre leur interprétation) et les économistes n'ont jusqu'à présent été que marginalement impliqués dans le projet (puisque les données n'étaient pas facilement exploitables sans connaissances relatives à l'histoire et à la méthodologie adoptée par TOFLIT18).

Néanmoins, ces difficultés ont eu le mérite de conduire à une collaboration multidisciplinaire intense, principalement entre historiens, historiens de l'économie et spécialistes des données. Cela permit de travailler avec le Medialab de Sciences Po et fut extrêmement enrichissant, malgré une substitution du temps en faveur des discussions d'ordre méthodologique et en défaveur du programme d'interprétation et, in fine, de la recherche.

Le principal objectif de TOFLIT18 était d'ouvrir une nouvelle voie pour la recherche sur l'économie française à la veille de la révolution industrielle grâce à l'analyse approfondie des données du commerce extérieur français. L'équipe de TOFLIT18 est fière de présenter à la communauté scientifique les outils de recherche produits dans le cadre de ce projet, rendu possible grâce à la collaboration entre historiens, historiens économiques et scientifiques et est impatiente d'achever le programme d'interprétation.

Guillaume Daudin

Références

Charles, L., et G. Daudin. 2011. « La collecte du chiffre au XVIIIe siècle: le Bureau de la balance du commerce et la production des données sur le commerce extérieur de la France ». *Revue d'histoire moderne et contemporaine* 58 (1): 128–155.

Charles, Loïc, et Guillaume Daudin, éd. 2015. « Eighteenth Century International Trade Statistics: Sources and Methods ». *Revue de l'OFCE (Special Issue)*, n° 140: 7-377.

O'Rourke, Kevin. 2017. « Review of Eighteenth-Century International Trade Statistics: Sources and Methods ». *The International Journal of Maritime History* 29 (2): 439-40.

Rössner, Philipp Robinson. 2008. *Scottish Trade in the Wake of Union (1700-1760): The Rise of a Warehouse Economy*. Stuttgart: Franz Steiner Verlag.