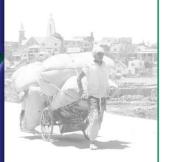## Focus :

Coming back to TOFLIT18

TOFLIT18 aims at opening new alley of research on the French economy on the eve of the Industrial Revolution by studying data on more than 500,000 French external trade flows extracted from the archives of the *Bureau de la Balance du Commerce* (L. Charles et Daudin 2011). We have already discussed our objectives in Dialogue #43 in July 2016. Three years after, we want to provide here a brief summary of how it evolved and where it stands now. More details can be found at https://toflit18.hypotheses.org. I believe our experience can be useful to other people starting similar endeavours.

First of all, the project was to enhance our knowledge of the structure and evolution of French local and national specialization. Second, the project was to focus on the interplay between political events and international trade, looking at the effect of the evolution of mercantilist economic policies, the effect of the *Exclusif* that reserved trade with French colonies to French ships, and the effect of wars on French trade and French economy. The quantitative data produced by the *Bureau de la Balance du Commerce* (L. Charles et Daudin 2011) is useful for all these aims because, though part of the data are macroeconomic by nature, a significant part of the local data are still available in the French regional Archives (e.g. Bayonne, Bordeaux, La Rochelle, Marseille, Nantes, Rouen…) and can be used to better understand the economic linkage between local economies, international trade and the national economy. Thus, analysing these data allow us to deepen our knowledge of economic phenomena on economic development and transformation and the determinants of international trade.

That requires collecting the existing data and making it useable for research. The data need to be worked with to ensure that they contain enough reliable information to be used safely as a guide to the past. Among the issues to be ironed out are their reliability, the difficulties of fitting the realities of merchant activities into administrative categories, the creation of reliable and stable list of goods and geographical entities ready for research and the treatment of prices.

Our research program addresses a crucial period, one that laid the economic ground for the entry of France and Europe in the modern era through the Industrial Revolution. International trade is a very important piece of the 18th century economic history puzzle. Whereas extra-European trade is comparatively well known and has been the object of recent synthesis, intra-European trade has been somewhat neglected. Only a few studies, all partials and based on heterogeneous methods, exist on this subject. More generally, their results are often too general and fragile to support precise economic analysis and comparisons.

The scarcity of works on foreign trade statistics of the early-modern period is all the more unfortunate as external trade flows have been the single economic data that early modern states have collected with the most care. Indeed, the first attempts at measuring foreign trade can be dated from the early seventeenth century. Annual series have been located for Sweden from 1732, Austrian Netherlands from 1759, Milan in 1762, 1769, 1778 and 1791, Venice from 1769/1770 onward, Portugal in 1776-77, 1783, 1789, and from 1796 onward, Spain from 1778 onward, etc. Research on trade of the British Isles is more advanced than for other countries. A recent book has renewed the research on Scottish trade in the 18th century (P. R. Rössner 2008). A team coordinated by David Jacks, Kevin O'Rourke, Alan Taylor is currently gathering data on English and British trade from 1696 to the early 20th century.

In the French case, the breadth of the data collected by the French state as well as their complexity made them more difficult to handle. The dispersion of a large part of the sources during the French Revolution increased the difficulty of their exploitation for economic historians. For this reason, most studies on "national" French external trade in the 18th century have used a limited range of archival sources. Country-level studies never used the whole range of centrally-gathered French data. A wealth of works of international trade and shipping in France at the local level, especially ports, include reflections on the economic effect of external trade, but they do not provide an in-depth discussion at the country level. Recent research works have shown more interest for this issue and some bilateral trade between France and other countries/European regions have been closely examined. These works do not however develop a systematic quantitative treatment of total French trade and do not use econometric tools.

1. How did we do it?

The first condition for the project success was the quality and the organization of the team. We operated on the basis of monthly internal meetings and annual conferences. Regular meetings allowed the discussion of critical issues, the development of common methodologies and the deepening of mutual understanding. Each year, our work was presented to the scientific community to trigger general discussions on methodological issues or interpretative research.

The prerequisite for the core of TOFLIT18 was the constitution of a database of the French external trade registered by the *Bureau de la balance du commerce* from 1716 to 1821 (including local data from 1716 to 1781). Guillaume Daudin and Loïc Charles already had the experience of *Bureau de la Balance du Commerce* data collection and re-transcription. The group included members of other database projects. The conceptual issues were simple enough as the set of raw data is relatively coherent. The main obstacle was the amount of work required for transcription, accuracy checks and dealing with the data flows.

A serious difficulty is linked to the classification of merchandises. Researchers need a general table of correspondence between the denomination of goods in different time periods and different places. Furthermore, they need goods to be identified and categorized to process and analyze the data. We had to realize a common user interface to organize the data that could accommodate multiple potential approaches of researchers. It was to be flexible enough that it can be re-used in the future to answer other research questions concerned with the material world of the eighteenth century and early nineteenth century. Building on past experience of the FNSP/Sciences Po médialiab, we selected the "datascape" methodology to answer these objectives. A "datascape" is a set of digital data methods and tools that fosters the use of exploratory data analysis (Tukey 1977) for social sciences research purposes (Latour and al. 2011). It aims at providing an exploratory data analysis method to model information from archives through navigation through archives. In our context, we used a bottom-up process for treating and classifying the data in categories. In short, we digitized the documents in a form (including spelling) as close as possible to the original and then, step by step, we normalized the data in order to supress (for example) spelling or geographical diversity without losing information. This provides a solid base to create classifications that relates to specific research questions. For

example, to study the international market for drug it was necessary to identify the commodity names that – in an eighteenth-century context – describe drugs and separate them from other kind of commodities. E.g. should mercury be treated as a medical drug? To do such thing, we had to develop expert knowledges through the use of eighteenth-century dictionaries and encyclopedias mostly.

2. Results

The first result is the creation of a unique database of French 18th century international trade that currently gathers nearly 515,000 verified trade flows to which will be added tens of thousands of additional ones still to be verified. The documented csv files are kept in a private GitHub repository (https://github.com/medialab/toflit18_data) that will be made public as soon as the datapackage documentation is finished (normally in 2019). Though anyone who wants to access it can ask us and be given access.

The completion of the datascape is certainly the most impressive achievement of TOFLIT18 (http://toflit18.medialab.sciences-po.fr/#/home). This is actually a formidable tool to explore a large quantity of heterogeneous data. The product glossary explores the meaning of more than 2,000 different obscure goods. The goods classifications reduce more than 55,000 goods denomination in than 20,000 ones without any information loss, and makes available a 25 category classification inspired from SITC that is a sound starting point for researchers. The datascape offers three different views on French trade (Time series, Product terms network, Locations network), and two additional ones for the French data (Metadata and Classifications). All the code is available under an open access license on https://github.com/medialab/toflit18.

These two major achievements have necessitated a lot of time resources that we originally intended to devote to analytical work. However, a first work has already been published and other ones are about to be published. The collective issue on comparative international trade statistics (Loïc Charles et Daudin 2015) has been called "an essential reference work for anyone interested in early modern trade" (O'Rourke 2017). It is available in free access (https://www.ofce.sciences-po.fr/publications/srevue.php?num=140).

A work on the economic consequences of the French wars is in its final stage and should be published in a collective volume directed by Patrick O'Brien (LSE). Still, a number of comparative works have been published or are at working paper stage. The comparison of the French data with comparable sources from the Sound strait, from Hamburg and from Belgium is well advanced and shows that they are sound basis for historical and economic work as long as some caution is exercised. In particular, the final version of the Sound paper is written and will be published soon in a collective volume in English. First version of papers presenting the datascape and the way the classification of goods evolved in French statistics have already been presented in seminars and conferences. We are preparing a French version of this methodological piece which is to be published in the *Revue de Synthèse* in early 2019. Eventually, works on drug trade, the effect of the loss of Canada, the effect of the Second Hundred Year War, and the functioning of grain trade have been started.

As a side note, we want to mention two achievements that we did not foresee when we put the project together. First, we have introduced to research nearly twenty interns from all levels of qualification (L3 to M2). Some of them have decided to continue research careers (including one in economic history). Second, we have fully documented the progress of the project on an Hypotheses blog (https://toflit18.hypotheses.org). We believe this additional work was worth it notably because we have put online the minutes of our monthly meetings. As a project manager, I know I would have been happy to access this kind of information beforehand to realize the type of difficulties I was to be confronted with.

TOFLIT18 created a datascape that is a real research infrastructure. It is the stepping stone to the creation of an European-wide network of similar research infrastructures.

3. What did we learn?

We obtained only 72% of the money we asked for. We are grateful for the money we had. Still, it led us to a number of decisions that slowed process of acquiring, digitalizing and curating the data. Because they were too expensive, we could not rely on professional transcribers as much as we could have done and had to delegate the bulk of the work to interns. It was difficult to find interns. It was difficult to manage them (the team leaders could not manage more than 4 interns at a time). They worked well enough, but their output was of heterogenous quality and required thorough verification.

This last issue highlighted the high price of not being able (for budgetary reasons) to hire a permanent and well qualified data administrator (ingénieur d'étude) for a year and a half as we had first planned to do. He or she would have had to be versed enough in paleography to be able to check the transcriptions and tech-savvy (although this last part could have been part of the training we may have offered) enough to manage the data flows from original archives to datascape-ready data through pictures, raw transcriptions and corrected transcriptions. The budget restriction of about 100,000€ was almost the exact full cost of such a qualified data administrator and made it impossible to hire such a person.

As a result, most of the curation work was done by one of the team leaders – a full time professor and researcher – helped occasionally by former interns who were hired for short periods (from 1 to 6 months). Inevitably, it was not done as quickly as we had hoped (and it is still not finished). This had two consequences. First, the interpretative program was much delayed as one needs to have the "full" range of data before starting the interpretative work. Second, the economists bailed out. One explanation is that the data were not in an easy-enough format to work without delving into historical-methodological issues for which they felt unqualified.

A silver lining in that cloud is that the intra-disciplinary work was done mainly between historians, economic historians and data scientists. Personally, I found that very enriching and I am retrospectively very happy to have worked with the Medialab in SciencesPo, but on the other hand the time passed to discuss methodological issues was unavailable for pure research which, again, delayed the interpretative program.

The main objective of TOFLIT18 was to open a new alley of research on the French economy on the eve of the Industrial Revolution by subjecting French external trade data series to an in-depth study. We are proud of the research tools we produced for the community thanks to the close collaboration between historians, economic historians and scientists. We look forward toward completing the interpretative program.

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*. Stuggart: Franz Steiner Verlag.