

April 2017

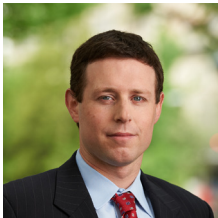


Dr. George Korenko

Partner | Washington, DC

gkorenko@edgewortheconomics.com

+1 202.559.4408



Mr. Matthew Milner

Partner | Washington, DC

mmilner@edgewortheconomics.com

+1 202.559.4387

Best Practices For Managing Data In Antitrust Cases

In antitrust litigation, economic experts rely on transactional sales data to study class certification issues, assess liability and calculate economic damages. Collecting these data, understanding how to interpret them and assembling them into a structure amenable to expert analysis is critical in formulating and testing empirical models and evidence. However, company data are often voluminous in size and spread across corporate entities and electronic systems around the world, making their collection and assessment complex. Efficient and effective collection and preparation of these data for analysis in litigation requires careful thought and planning. Throughout this process, attorneys and experts should work with their clients to address the following questions about transactional data:

- What systems contain potentially relevant transactional sales data?
- What is the best way to extract the relevant data for litigation?
- How can the data be validated?
- What steps should be taken to prepare the data for analysis?

In this article, we address the questions above and discuss best practices for lawyers and experts for the efficient and effective collection and preparation of transactional data in antitrust cases. Through the article, we apply the concepts we discuss to the experience of two hypothetical companies, “Toolmaker” and “Tokyo Tooling,” which manufacture and sell tools to customers in Japan and the United States.

Hypothetical Antitrust Case

As happens in a typical case involving claims of price-fixing, suppose that Toolmaker Asia and its subsidiary, Toolmaker USA, have been accused by a putative class of direct purchasers of conspiring with Tokyo Tooling to coordinate

the prices of screwdrivers sold in the United States. As part of discovery, the plaintiffs' "Request for Production of Documents" includes the following:

Request for Production No. 1

Your transaction-level sales data for all worldwide sales of screwdrivers from January 2006 to December 2016, including:

- Date of transaction, including date of order, date of invoice, and date of shipment.
- Invoiced unit price and currency in which it is expressed.
- Quantity sold (including negative values that reflect returns or credits).
- Total amount of the transaction, and currency in which the amount is expressed.
- Any discount or surcharge associated with each transaction.
- Product characteristics, including product number to uniquely identify the screwdriver product, brand, product line, blade type, and grip.
- Detailed customer information, including bill-to and ship-to customers and addresses.
- Detailed costs associated with sales of all screwdriver products.

Responding to this request can be demanding and time-consuming. However, by adhering to certain best practices for the collection of data, outside counsel and experts can assist the company with responding to the plaintiffs' request as efficiently as possible.

What Systems at Toolmaker Contain Potentially Relevant Transactional Sales Data?

Discovery in antitrust litigation often involves collection of vast amounts of data, which can vary in type, the database systems in which they are contained, as well as where in the world they are located. In many cases, the data may also be stored in different languages. Given the array of data typically requested and the numerous systems involved, the task of identifying the relevant sources can be daunting for companies. A key to identifying relevant systems is to understand how the company maintains data and information.

In this case, the first step in collecting data for Toolmaker is to understand what transactional data are available. As a corporation doing business around the world from several international sales offices, the company's transactional sales data are maintained in disparate systems across countries and subsidiaries. For example, Toolmaker Asia uses an enterprise resource planning ("ERP") system (e.g., Oracle or SAP) to maintain sales, accounting, human resources, manufacturing and inventory data. However, Toolmaker USA maintains its data in an in-house system and stores most of its historical records on a decommissioned mainframe and backup tapes. Armed with this knowledge about the pertinent information systems and technology, counsel, experts and the company will be able to develop a profile of potentially relevant data for this antitrust litigation. This profile will show data that are actively used by the company, files that are archived, and the scope of the company's network containing potentially relevant information.

Toolmaker's business people are knowledgeable about the company's current data systems. However, given the 11-year period covered by the discovery request, there are few people at Toolmaker USA that are knowledgeable about the legacy systems containing the historical data. This will be an important issue to consider when assessing

the historical data relevant for the case and the burden and the resources needed to collect and extract these files. Because there are few people at the company that are knowledgeable of the decommissioned systems, other sources of information will be important to make an assessment. For example, Toolmaker kept documentation on the transition from its legacy system to its current system, including the data tables and fields that were archived, and information that was transferred to the current system.

Documentation on the legacy system at Toolmaker will be instructive to determine if the files are accessible and contain relevant information, which will help shape the plan to restore the data.

What Is the Best Way to Extract the Relevant Data From Toolmaker?

Toolmaker has maintained a substantial volume of sales and cost data and information, but not all this information may necessarily be relevant for analyzing an antitrust claim. Toolmaker Asia’s ERP system has thousands of data tables, but most of this information is not relevant for analyzing price fixing allegations. For example, the human resources records that contain personally identifiable information do not pertain to the antitrust allegations and could raise data privacy concerns. Rather, a targeted plan for identifying and extracting only tables relevant to pricing — rather than request a “data dump” of the entire ERP system — would be more effective. This approach focuses on the relevant sales and pricing information on a customer invoice to identify the transactional sales data requested.


An often-effective approach for targeting relevant pricing data within vast enterprise systems is to collect an actual paper invoice and identify the key data elements represented on it. For example, consider the invoice below, which Toolmaker Asia issues to its customer Screw Drivers R' Us.

ORIGINAL INVOICE

TOOLMAKER ASIA LTD.
52 Nagaïke-cho,
Yokohama-shi, Kanagawa 235-8534, Japan

BILL TO:
Screw Drivers R' Us
Att: Accounts Payable
PO Box 435
Witchita, KS 67203

SHIP TO:
Screw Drivers R' Us
3200 Industrial Blvd
Tulsa, OK 74101



Refer to this number on correspondence

INVOICE NUMBER	100135
INVOICE DATE	25JAN2010
SALES REP	TAYLOR

DIRECT ALL INQUIRIES TO:
Akira Saito

PO NUMBER	ORDER DATE	WAREHOUSE	SHIP DATE	BILL OF LADING	SHIPPING	TERMS
54355	23JAN2015	120	02FEB015	5630021	AIR	2/15 NET 30

QUANTITY		UOM	UPC	Line	ITEM DESCRIPTION	UNIT PRICE	ALLOWANCE	NET UNIT PRICE	EXTENDED AMOUNT
ORDERED	SHIPPED								
200	200	CA	68158	1	TMA No. 3x6 in Srewdriver, Phillips	¥ 533	¥ 35	¥ 498	¥ 99,600
30	15	CA	43291	2	TMA Prem 17 pc. Screwdriver Set	¥ 1,423	¥ 100	¥ 1,323	¥ 19,845
10	10	CA	12999	3	TMA #00 x 1-1/2 in Screwdriver	¥ 340	¥ -	¥ 340	¥ 3,400
8	8	PAL	99111	4	TMA23Z 5 pc Torx SD Set	¥ 850	¥ 50	¥ 800	¥ 6,400
5	5	CA	23452	5	TMA44SZ 1/4 in. Nut Driver	¥ 584	¥ -	¥ 584	¥ 2,920

TERMS 2/12 NET 30 If paid by: 02MAR015 MM/DD/YY Discount of: \$ ¥ 2,643	SUBTOTAL ¥ 132,165 TAX ¥ 6,608 FREIGHT ¥ 3,965 GRAND TOTAL ¥ 142,738 Pay above amount if discount is not earned.
--	--


The invoice shows that Toolmaker sold five different tool products to Screw Drivers R' Us on Jan. 25, 2010. The invoice also contains information on each of the products sold, including the UPC codes, unit prices, allowances, sales quantities, and extended amounts. Specifically, we see that the company sent customers detailed information on pricing ("unit price"), products ("UPC" and "item description"), discounts ("allowance"), and sales ("quantity shipped" and "extended amount"). There is also information about the invoice number, order number and dates associated with the order, shipment, and invoice. Here, the total amount invoiced to Screw Drivers R' Us amounted to ¥142,738.

This invoice from Toolmaker Asia contains information that is responsive to the plaintiffs' request and is useful for analysis by experts. The key insight here is that the invoice is generated by Toolmaker's sales system. Identifying where each of the data elements on the invoice resides in the system can serve as a roadmap for extracting the relevant pricing data. The invoice provides a roadmap to identify the relevant sales and pricing data from the company's systems.

Through discussions with Toolmaker's database management personnel, and with this invoice in hand, counsel and the expert economist can work to identify the relevant portions of the sales system where these data reside. For example, the annotated invoice above breaks the invoice that the customer receives up into the locations of these data elements in Toolmaker's systems. For example, the bill-to and ship-to customer information are maintained in the "customer address" data table. Similarly, product information, shipping information and pricing information are all tracked in specific electronic tables maintained by the company. Rather than searching through hundreds or thousands of data tables and fields across an ERP system, this process narrows the scope to key tables and fields with sales, product, and pricing information. Working with the company to map the systems and tables to the invoice facilitates tracking the prices and associated information sent to customers. Further, by identifying the data tables and fields where the information on the invoice is electronically stored, counsel for Toolmaker can ensure that relevant data are extracted.

ORIGINAL INVOICE

TOOLMAKER ASIA LTD.
52 Nagaike-cho,
Yokohama-shi, Kanagawa 235-8534, Japan



BILL TO:
Screw Drivers R' Us
Att: Accounts Payable
PO Box 435
Witchita, KS 67203

SHIP TO:
Screw Drivers R' Us
3200 Industrial Blvd
Tulsa, OK 74101

MAKE ALL CHECKS PAYABLE TO:
TOOLMAKER ASIA LTD
Attn: Accounts Receivable
52 Nagaike-cho,
Yokohama-shi, Kanagawa 235-8534
Japan

Refer to this number on correspondence

INVOICE NUMBER 100135
INVOICE DATE 25JAN2010
SALES REP TAYLOR

DIRECT ALL INQUIRIES TO:
Akira Saito

PO NUMBER	ORDER DATE	WAREHOUSE	SHIP DATE	BILL OF LADING	SHIPPING	TERMS
54355	23JAN2015	120	02FEB015	5630021	AIR	2/15 NET 30

QUANTITY		UOM	UPC	Line	ITEM DESCRIPTION	UNIT PRICE	ALLOWANCE	NET UNIT PRICE	EXTENDED AMOUNT
ORDERED	SHIPPED								
200	200	CA	68158	1	TMA No. 3x6 in Screwdriver, Phillips	¥ 533	¥ 35	¥ 498	¥ 99,600
30	15	CA	43291	2	TMA Prem 17 pc. Screwdriver Set	¥ 1,423	¥ 100	¥ 1,323	¥ 19,845
10	10	CA	12999	3	TMA #00 x 1-1/2 in Screwdriver	¥ 340	¥ -	¥ 340	¥ 3,400
8	8	PAL	99111	4	TMA23Z 5 pc Torx SD Set	¥ 850	¥ 50	¥ 800	¥ 6,400
5	5	CA	23452	5	TMA44SZ 1/4 in. Nut Driver	¥ 584	¥ -	¥ 584	¥ 2,920

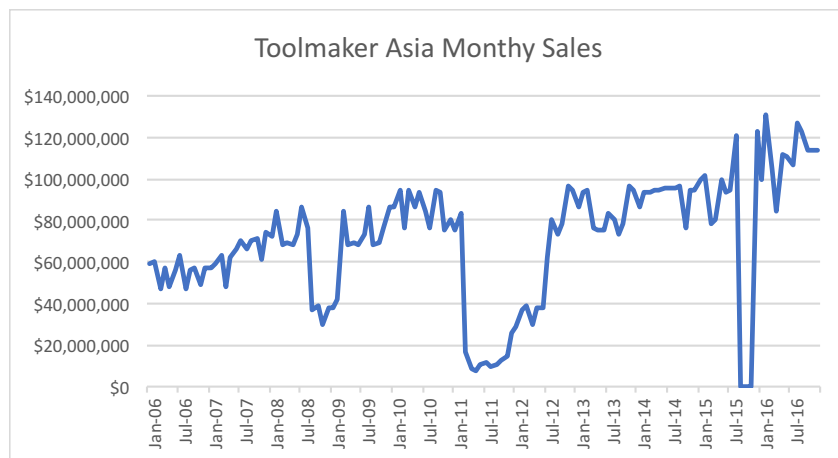
TERMS	SUBTOTAL	¥
2/12 NET 30		132,165
If paid by: 02MAR015	TAX	6,608
MM/DD/YY	FREIGHT	3,965
Discount of: \$ ¥ 2,643	GRAND TOTAL	¥ 142,738
	Pay above amount if discount is not earned.	

The invoice to ERP mapping separates relevant data elements from the many that are not needed for the litigation. With the relevant tables and fields identified, the next step is to develop an efficient process to extract the actual files from Toolmaker. Exporting historical data is often an iterative process and may require coding by company staff and use of company computing resources. To streamline this process, a best practice is to export a sample of the data from a shorter period, which consumes less of the client’s time and resources. The small sample extraction from Toolmaker systems allows counsel to ensure the correct elements mapped out from the invoice are exported, assess how the records in the data are populated, and determine if there were any issues with the files during the export (e.g., corrupted files or duplicate records). This careful assessment of sample data mitigates the risks of data integrity issues during the full data extraction — the goal being that the full data are extracted once, saving the client time and resources.

How Can the Toolmaker Data be Validated?

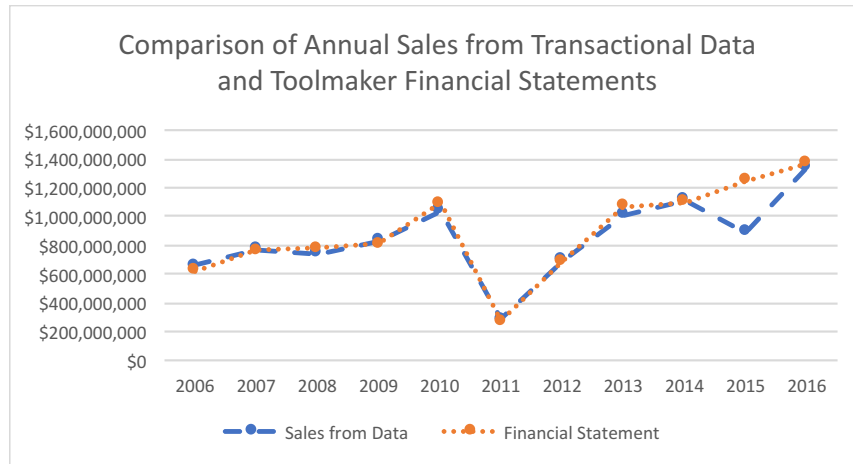
When the historical data have been exported from Toolmaker’s systems, there are certain techniques that can be employed to ensure the data are free of duplicates and omissions. These types of issues can occur during the extraction process because data that is kept during the ordinary course of business is not pristine — data entry errors, system updates and outages can result in data errors. A best practice is to validate the data using external sources such as financial statements and studying patterns in the data, which can help identify potential issues with the files. For example, preparing monthly sales reports and transactional counts can show if there are potential data omissions. Substantial changes over time in either the total volume of commerce or the number of transactions may raise questions — sudden drops in the data may be due to omissions or sudden spikes in the data can be due to duplicate transactions.

For example, the graph below show Toolmaker Asia’s sales between 2006 and 2016, using an initial data extract from Toolmaker’s systems.



Notably, there is a decline in late 2008, and a more dramatic decline in sales in 2011. In addition, there appears to be a complete absence of sales at the end of 2015. To investigate these potential anomalies, the transactional sales data are aggregated and compared to Toolmaker Asia’s annual financial statements. Small differences between annual

sales amounts may be explained by differences in the timing of invoicing and the recognition of sales transactions for accounting purposes. However, if the sales amounts derived from the transactional data and the financial statements are materially different, there may be omissions in the sales data that warrant further exploration.



The graph above confirms that the annual sales derived from Toolmaker Asia’s transactional data closely match the sales in the company’s financial statements for all years, except 2015. Consultations with Toolmaker’s business personnel can confirm that the decline in 2009 is explained by the financial crisis and the steep decline in 2011 was a result of a tsunami that damaged Toolmaker Asia’s plant. The analysis and discussions with the business people at Toolmaker can be used to confirm that the gap in 2015 was due to omitted transactional data. This type of validation analysis can help distinguish between actual data quality problems (which could create issues in terms of complying with the plaintiffs’ request) and real-life fluctuations in the business over time.

What Is the Best Way to Prepare the Data for Expert Analysis?

Once the transactional data from Toolmaker and co-defendant Tokyo Tooling are collected and validated, the next step is to prepare the files for analysis, often referred to as the “data build.” This stage could involve processing hundreds of files and terabytes of data. There is no standard template for data preparation; rather, it is a process where the expert gains knowledge about the company’s business and conducts empirical analyses of the alleged conduct.

Best practices in building data for expert analysis include careful thought about: (1) identifying and understanding key elements in the data; (2) combining disparate data sources; and (3) additional information that should be incorporated with the sales data.

Identifying and Understanding Key Elements in the Data

When data are produced for litigation it is important to ask questions about the elements observed in data, but, more importantly, key in on the questions that are most relevant to the case. In this price-fixing case, the nature of screwdriver pricing and, specifically, how the alleged conduct affected pricing of different products and to different customers are key elements. Fields that contain information associated with these elements should be the focus of the data build. For example, a detailed inquiry into the price fields for Toolmaker and Tokyo Tooling reveals that

Toolmaker's price field includes the cost of shipping, but Tokyo Tooling's does not. Time spent exploring key elements such as pricing will influence how the data are prepared and analyzed.

Decisions on Combining Disparate Data Sources

Before the data are processed for each company and system, careful thought should be given to what fields and elements are similar across data sources. For Toolmaker and Tokyo Tooling, we have disparate datasets in English and Japanese, with differing time periods and numerous fields that do not correspond across datasets. Even seemingly similar data fields may contain information that is not directly comparable. For example, since Toolmaker's price field includes the cost of shipping, but Tokyo Tooling's does not, combining the two data sets into a single "price" field will result in comparisons across the defendants that are not apples-to-apples. In this case, it would be more useful to maintain separate databases for Toolmaker and Tokyo Tooling.

Considerations for Incorporating Additional Information

There are often other sources that can be used in conjunction with transactional sales data to create a data set with richer information. In this case, Toolmaker and Tokyo Tooling have additional files that can be linked with their respective sales records. These include:

- Product decoders, which provide more detailed information on screwdriver characteristics associated with each unique product number.
- Customer detail lists, which provide additional information from the marketing department on customer type (e.g., mass merchandiser, hardware store, wholesaler).
- Cost information from other sources may be mapped to prices at a product level. This includes standard costs for materials, processing, and labor.

The decision to map the above information to the transactional data will depend on the availability and usefulness of the information, the overlapping periods with the sales data, and if it is appropriate to map such information on a transaction-by-transaction basis.

Conclusion

The collection and preparation of transactional data is critical to a successful economic analysis in an antitrust case. Best practices involve collaboration between the client, counsel and expert to ensure that the relevant data are collected and prepared in an efficient and effective manner. This includes:

- Identifying the data systems that potentially contain data requested in the discovery phase of the case;
- Carefully considering the best way to extract the relevant data, including assessing whether a targeted approach may be more efficient and effective than a full data dump;
- Validating the extracted data to identify potential anomalies;
- Thoughtfully assessing how to prepare the data for analysis, including identifying key elements in the data, determining the extent to which disparate datasets should be combined, and evaluating whether there is additional useful information that can be linked together.

Following these best practices is likely to yield data that can be reliably used to address relevant economic questions posed by counsel and experts. ■

About Edgeworth

Edgeworth Economics provides quantitative and economic consulting in the course of litigation and business to its clients, which include world-class law firms, Fortune 500 companies, and government agencies. Edgeworth experts apply their knowledge and experience, along with state-of-the-art computing infrastructure, to help clients efficiently manage complex issues including antitrust litigation, consumer financial services, privacy & data security, transfer pricing, intellectual property, mergers and acquisitions, class actions, labor, and data & HR analytics. As a rapidly growing firm with a fresh approach, Edgeworth attracts leaders and teachers from across the industry including PhD economists, MBAs, statisticians, and programmers. Edgeworth has offices in Washington, DC, Pasadena, and San Francisco.

www.edgewortheconomics.com

Disclaimer:

The opinions expressed herein do not necessarily represent the views of Edgeworth Economics or any other Edgeworth consultant. This article is intended to inform readers about legal developments. Nothing in this article should be construed as legal advice or a legal opinion, and readers should not act upon the information contained in this article without seeking the advice of legal counsel.