Improving Inbound Visibility through Shipment Arrival Modeling

by

Michael Chun

B.S. Operations Research, United States Military Academy, 2007

Submitted to the MIT Sloan School of Management and the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Engineering Systems

In conjunction with the Leaders for Global Operations Program at the Massachusetts Institute of Technology

June 2014

© 2014 Michael Chun. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature redacted

Signature of Author

Engineering Systems Division, MIT Sloan School of Management

Certified by

Edgar Blanco, Thesis Supervisor
Research Director, MIT Center for Transportation

Certified by

Karen Zheng, Thesis Supervisor
Assistant Professor of Operations Management

Accepted by

Richard C. Larson, Mitsui Professor of Engineering Systems
Chair, Engineering Systems Division Education Committee

Accepted by

Maura Herson, Director of MIT Sloan MBA Program
MIT Sloan School of Management
This page intentionally left blank
Improving Inbound Visibility through Shipment Arrival Modeling

by

Michael Chun

Submitted to the MIT Sloan School of Management
and the Engineering Systems Division on May 9, 2014
in Partial Fulfillment of the Requirements for the Degrees of Master of Business Administration
and Master of Science in Engineering Systems

Abstract

Amazon’s desire to provide the “Earth’s largest selection” comes at a tremendous cost. In addition to managing the orders for thousands of vendors/suppliers and millions of Stock Keeping Units (SKUs), Amazon must keep track of all the inbound shipments in order to manage its inventory efficiently. Although estimated delivery dates are routinely received for each of these inbound shipments, only about half of the purchase orders actually arrives by these dates. Since knowing exactly how much to order is based in part by what has already been ordered in the past and when those shipments will arrive, this inaccuracy makes determining optimal purchase order quantities difficult for future orders.

So in order to optimize the inbound process, Amazon must either improve the accuracy of these estimates or account for the inherent variation. This thesis establishes a model that exposes the underlying variation for each inbound arrival signal based on historical error rates. Our approach is to map all inbound signal sources and then create a classification-tree model that minimizes the joint variance of the prediction errors. Simulations indicate that such a model can be used to generate new estimated arrival dates that reflect the likelihood of arrival. In addition, this thesis also takes a step further to outline some potential vendor policy changes for eliminating the root causes of procurement lead time variance.

Thesis Advisor: Edgar Blanco
Research Director, MIT Center for Transportation

Thesis Advisor: Karen Zheng
Assistant Professor of Operations Management, Sloan School of Management
This page intentionally left blank
Acknowledgements

I would like to begin by thanking both Amazon.com and MIT’s Leaders for Global Operations (LGO) program for the incredible opportunity to learn and contribute to a meaningful project. Special thanks to Dan Joy, David Greenstein, Maneesh Jyoti, Robert Gash, and Stephen Brown for their mentorship, insight, time, and support throughout my internship.

Also, special thanks to my thesis advisors, Professor Edgar Blanco and Karen Zheng, for providing guidance on tackling complex problems and applying theoretical knowledge to real-world challenges.

Since day one, I really felt like an Amazonian thanks to the amazing and supportive teammates of the Inbound Supply Chain Team. So thank you Molly Boatright, Andrea Lau, Kristin Lien, Shruti Mathure, Bijal Mehta, Joseph Meyer, Samyukta Soothram, Li Yuepeng, and the rest of the team for helping me with everything from SQL queries to deciphering the endless list of acronyms.

This internship would not have been possible without the time and support of all of the great software developers who helped me to understand and improve the complex IT infrastructure driving Amazon’s operations. Thank you Gary Zhao, Rob MacGregor, David Robson, Chris Chamberlin, Samad Jawaid, Jordan Foo, and Bryce Moore.

Finally and most importantly, I would like to thank my loving wife Jamie and daughter Audrey for joining me on this journey and giving me a reason to smile each and every day. After all, it was Jamie’s encouragement and insight that led to me LGO to begin with. Thank you for everything—I couldn’t have done it without you.
This page intentionally left blank
**Table of Contents**

Abstract .......................................................................................................................... 3

Acknowledgements ........................................................................................................ 5

Table of Contents ......................................................................................................... 7

List of Figures ............................................................................................................... 9

1. Introduction and Background .................................................................................. 12
   1.1 Purpose .................................................................................................................. 12
   1.2 Amazon.com .......................................................................................................... 12
   1.3 The Inbound Supply Chain Challenge .................................................................... 14
   1.4 Defining the Problem ............................................................................................ 17
   1.5 Thesis Overview ..................................................................................................... 18

2. Literature Review ....................................................................................................... 20
   2.1 Data Mining Approaches .................................................................................... 20
   2.2 Divide-and-Conquer with Classification Trees ................................................... 21
   2.3 Attribute Selection ............................................................................................... 23

3. The Inbound Process .................................................................................................. 25
   3.1 An Overview .......................................................................................................... 25
   3.2 The Procurement Process .................................................................................. 26
   3.3 The Inbound Transportation Process .................................................................. 28
   3.4 The Inbound FC Process .................................................................................... 29

4. Data Collection, Analysis, and Model Development ............................................... 30
   4.1 Data Collection from Potential Signal Sources .................................................. 30
   4.2 Attribute Selection for Model Development ..................................................... 32
   4.4 Distribution Selection ......................................................................................... 36
   4.5 Using the Classification Tree ............................................................................... 37

5. Classification Tree Model Validation ....................................................................... 38
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Testing Parameters</td>
<td>38</td>
</tr>
<tr>
<td>5.2 Simulation Assumptions and Limitations</td>
<td>39</td>
</tr>
<tr>
<td>5.3 Simulation Results</td>
<td>40</td>
</tr>
<tr>
<td>5.4 Using the Proposed Classification Tree Model</td>
<td>42</td>
</tr>
<tr>
<td>5.5 Simulation Conclusions</td>
<td>43</td>
</tr>
<tr>
<td>6. Beyond Tracking Inbound Signals</td>
<td>44</td>
</tr>
<tr>
<td>6.1 The Role of Vendor Compliance</td>
<td>44</td>
</tr>
<tr>
<td>6.2 Supplier-Retailer Partnerships</td>
<td>45</td>
</tr>
<tr>
<td>6.3 The Path Ahead</td>
<td>46</td>
</tr>
<tr>
<td>Bibliography</td>
<td>48</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Amazon revenue over ten years.................................................................13
Figure 2. Rising shipping costs as a percentage of net sales.................................13
Figure 3. Amazon's growth strategy.......................................................................14
Figure 4. Histogram of the error rates for all inbound signals (sampled data from January to June 2013) .................................................................17
Figure 5. Exposing the underlying variability associated with each signal ............17
Figure 6. An overview of the Knowledge Discovery in Databases (KDD) process .................................................................21
Figure 7. General classification tree structure........................................................22
Figure 8. A two class example with independently and identically distributed (i.i.d.) variables.................................................................23
Figure 9. Amazon supply chain overview................................................................25
Figure 10. FC processes ......................................................................................25
Figure 11. Planning logic and key decisions............................................................26
Figure 12. The procurement process........................................................................27
Figure 13. Routing request process.........................................................................28
Figure 14. Inbound FC process ............................................................................29
Figure 15. Overview of the inbound process and key predictive signals....................31
Figure 16. Boxplot of prediction error by signal source ........................................32
Figure 17. Classification by signal source .............................................................32
Figure 18. Classification by signal source and lowest joint variance attribute ..........35
Figure 19. Sample terminal node distribution..........................................................35
Figure 20. Sample error distribution (Signal 2) .....................................................38
Figure 21. Sample arrival distribution with the shaded area representing 90% of the distribution .................................................................38
Figure 22. Base model and classification tree model prediction accuracy comparison ..................................................................................41
Figure 23. Average absolute error between arrival-by date and receive date ........................................42
1. Introduction and Background

1.1 Purpose

The purpose of this thesis is to build a model that accounts for the variance in procurement lead times associated with vendor shipments and deliveries. From the moment a purchase order is created, there are many signals that estimate the arrival date of the corresponding shipment. However, only 52% of purchase orders actually arrive by the estimated delivery date and only 21% of orders arrive on the estimated delivery date. Creating a more accurate way of assessing the arrival dates of inbound purchase orders is a critical step towards optimizing the inbound supply chain. This research was conducted during a six-month internship in partnership with Amazon.com and the Leaders for Global Operations program from the Massachusetts Institute of Technology.

1.2 Amazon.com

Founded in 1994 by Jeff Bezos, Amazon.com dedicates itself “to be Earth’s most customer-centric company, where customers can find and discover anything they might want to buy online, and endeavors to offer its customers the lowest possible prices.” From upending the publishing industry with its Kindle e-book devices to opening its web-retail platform to third party merchants, Amazon continues to find new and exciting ways to fulfill its mission and delight its customers every day.

Yet the core of Amazon’s business continues to be the sale of physical goods through its extensive network of warehouses referred to as Fulfillment Centers (FCs) around the world. By investing about $13.9 billion, Amazon has more than doubled the number of FCs to 89 from the year 2010 to 2012. Such rapid development of infrastructure is necessary to keep up with the rapidly growing annual sales:

---

1 “Amazon.com: About Amazon.”
2 “Why Amazon Is on a Warehouse Building Spree - Businessweek.”
In addition, positioning its warehouses closer to its customer base allows Amazon to achieve faster shipping times, while simultaneously reducing its outbound transportation costs—a particularly important element in the face of steadily rising shipping costs:

Figure 1. Amazon revenue over ten years

Figure 2. Rising shipping costs as a percentage of net sales
This rise in shipping cost is a direct result of both free-shipping offers and the Amazon Prime program. Amazon Prime was launched in 2005, offering free two-day shipping to its members on a large selection of items for a flat annual fee. Although offering free shipping and shifting these costs away from customers appear counter-intuitive at first glance, programs like Amazon Prime fit perfectly within the growth strategy that Jeff Bezos outlined on a paper napkin during its early stages:

Figure 3. Amazon’s growth strategy

Everything begins with the customer experience. Providing amazing customer experience will generate more traffic to the Amazon website, which will then attract new sellers and expand Amazon’s product offering. As Amazon grows, it can leverage economies of scale to reduce prices. This combination of reduced prices and increased selection will then continue to enhance the customer experience, further promoting faster growth in an unending virtuous cycle. Offering free shipping then is simply one way to enhance the customer experience.

1.3 The Inbound Supply Chain Challenge

But offering “Earth’s Biggest Selection” comes at a tremendous cost. Not only does Amazon have to manage orders for thousands of vendors/suppliers and millions of SKUs, but it must also find a place to store this inventory. As its selection continues to grow at a rapid pace, simply building new FCs to accommodate for additional inventory is an expensive and unsustainable option.

---

5 “Amazon.co.uk: It’s Still Day One.”
6 “Amazon.com—Earth’s Biggest Selection.”
One way to reduce the need for additional warehousing space is to lower the levels of safety stock through improved inbound visibility. We define inbound visibility as knowing precisely what items will arrive at the FC and when that item will be available for purchase by an Amazon customer. Although Amazon communicates its ship and delivery expectations to its vendors and suppliers in its purchase orders, knowing precisely when a shipment will arrive still remains a tremendous challenge. The following three reasons help explain the challenges in making accurate inbound arrival predictions:

1. **Unpredictable Variability** -- There are some elements of the lead time that cannot be accounted for at the time the purchase order is created. This kind of variability includes bad weather, backorders, cargo theft, and missed appointments.

2. **Broad Vendor/Supplier Base** -- Working with thousands of different vendors means working with a broad range of vendor capabilities. Some vendors have sophisticated warehousing systems and can maintain consistent lead times while other “mom and pop”-like vendors have a poor track record of delivery compliance. In terms of inbound visibility, this means that not all inbound estimates are equally reliable. Furthermore, the IT capabilities vary widely from vendor to vendor. Some vendors are capable of sending near-real time updates using Electronic Data Interchange (EDI) codes, while other vendors manually update the status of their shipments through Amazon’s online portal.

3. **Multiple Ship Modes and Ordering Volumes** -- Keeping millions of SKUs in stock means ordering in a wide range of quantities and volumes. Amazon’s purchase order sizes range from small packages sent via USPS to full shipping containers sent via rail. This wide range of ship modes for its inbound orders is another significant source of variation in terms of lead times as the processes associated with each transportation method correlates with a wide range of lead times.

Consequently, both cost and service is severely impacted, driving down efficiencies in order placement and sometimes even customer fulfillment. Without clear visibility on inbound shipments,
products are either received too early (increasing inventory holding costs unnecessarily) or not received on-time (increasing opportunities for Amazon to stock out and miss customer promised shipments). But overcoming this inbound challenge could yield the following immediate and significant benefits:

1. **Optimal Inventory Ordering** – Reorder quantities are based on the predicted demand over the lead time plus some safety stock. So when the vendor lead time is accurately determined, we can reduce the amount of safety stock ordered.

2. **Extended Inventory Availability** – Although out-of-stock items at an Amazon warehouse can still be available for purchase by customers, the delivery date promised to customers can extend much further than what Amazon is actually capable of doing. This is often due to the absence of reliable inbound shipment data. If we always know exactly what shipments were arriving and when, we can provide much shorter delivery promise dates to customers, increasing the likelihood that a customer will purchase an out-of-stock item. For example, if one SKU ran out-of-stock today and we know with 100% certainty that a restocking shipment will arrive tomorrow, Amazon could still promise customers a delivery in two-days by sending the package via overnight shipping.

3. **Accurate Customer Promises** – Even when Amazon does run out-of-stock for a certain item, having an accurate inbound visibility is key to offering its customers reliable estimates on when they can expect to receive their order. Such accurate communication with customers could reduce the likelihood that customers go elsewhere to find the out-of-stock item.

4. **Accurate FC Labor Planning** – Variations in the volume of shipments that arrive at each FC is a daily challenge for inbound FC labor planners. Unwanted variation leads directly to over or under staffing, which translates directly to voluntary time-off or mandatory overtime (both are costly consequences in terms of employee morale and warehousing labor expense). Improving inbound visibility will help identify the unexpected surges and reductions in inbound shipment volume.
1.4 Defining the Problem

Amazon already aggregates all of the signals related to inbound inventory from the moment a purchase order is created to the time the ordered item is received at an FC. Although each signal provides a specific estimated delivery date, only 21% of these estimates are accurate. As evidenced in the chart below, these estimates are subject to a wide range of variability:

![Histogram of prediction error](image)

**Figure 4. Histogram of the error rates for all inbound signals (sampled data from January to June 2013)**

Our goal then is to accurately capture the varying levels of reliability associated with each inbound signal and expose this variability by generating an arrival distribution for each item through a predictive model:

![Predictive model diagram](image)

**Figure 5. Exposing the underlying variability associated with each signal**

---

7 Note: Many of the y-axis labels in this thesis has been altered or removed in the interest of protecting Amazon.com's confidentiality.
Translating estimated arrival dates into arrival distributions will help drive many other IT systems with application ranging from inventory placement decisions (which FC to send an inbound item) to future purchasing decision (how much of an item to order). Furthermore, each of these systems has a different level of sensitivity to how accurate the predicted arrival date needs to be and the distribution makes it easier to adjust for it accordingly.

1.5 Thesis Overview

The general approach taken in this thesis has been broken down into four phases:

**Phase 1 – Process Mapping**

Because we are primarily interested in identifying and interpreting relevant signals, the first phase is to chart the entire inbound process and map the underlying IT infrastructure supporting it. A detailed process map would reveal potentially useful prediction signals. In addition, the mapping process is important for understanding the business decision that our proposed model would impact.

**Phase 2 – Data Collection, Analysis, and Model Development**

Our primary aim is to develop a model that best exposes the underlying distribution associated with each inbound signal. Our approach is to split the sample data set along different attributes to identify the attributes that reduces the joint variance of the model. We will also determine the best distribution fit for these errors. With these insights, we can then build a model for interpreting future signals.

**Phase 3 – Classification Tree Model Validation**

Next we need to validate the functionality of our proposed model. Since the primary output of the model are new estimated delivery dates for a wide range of arrival probabilities, we will need to measure the error between these dates and the actual arrival dates within the simulation. We will also discuss other models that may be done in the future to improve upon the results presented here.

**Phase 4 – Beyond Tracking Inbound Signals**

It is important to note that the model presented here attempts to only capture the existing variability associated with delivery performance, but does nothing to eliminate the root causes of such
variability. So in the final phase, we will look beyond the prediction algorithm to identify possible vendor policy changes that can be made to improve the overall on-time delivery performance.
2. Literature Review

The research methodology outlined in this thesis explores the primary processes and signals that characterize the inbound process. To model this process, we consolidate all of the available data and even convert some of the data into a format that can be analyzed using different analytical techniques. As a result, this approach is directly analogous to a data mining process and the research presented in this paper lays the groundwork for more advanced machine learning algorithms in the future. Once this initial analysis is complete, it will be relatively straightforward to directly apply additional data mining techniques without significant effort spent on identifying, preparing, consolidating, and transforming the data.

2.1 Data Mining Approaches

With the rapid development of computing technology and the falling cost of digital storage, our ability to collect data has begun to outstrip our ability to interpret this data in a meaningful way. The term ‘data mining’ came about in the 1990’s to describe an emerging field that used a combination of classical statistics, artificial intelligence, and machine learning to make sense of large troves of collected data. In this section, I provide an overview of the modern approaches to data mining.

In the broadest sense, the purpose of data mining is to find patterns in data using different algorithms. However, as seen in the diagram below, data mining is simply one step in the overall process of knowledge discovery that really begins with data collection, selection, preprocessing, and transformation. More importantly, data mining could only be considered a success if the patterns found are relevant and insightful enough to be incorporated into other models for further action. This is why a distinction is made between the resulting patterns from data mining and “knowledge.”

---

8 Holmes, Tweedale, and Jain, Data Mining.
9 Fayyad, Piatetsky-Shapiro, and Smyth, “From Data Mining to Knowledge Discovery in Databases.”
The goals of the data mining process are to either verify a hypothesis or to discover new patterns. The discovery goal can be further subdivided into one of two categories: prediction (finding patterns that can be used to predict future outcomes) or description (finding patterns that characterize a system in a way that can be easily understood by humans). To achieve these goals, three major data mining approaches have surfaced over the last few decades:

1. **Classification** is the process of mapping (classifying) a data item into one of several predefined classes.
2. **Clustering** is the process of grouping similar objects into one cluster while partitioning dissimilar objects into different clusters.
3. **Regression** is the process of mapping a data item to a real-valued prediction variable.

### 2.2 Divide-and-Conquer with Classification Trees

The approach taken in this thesis falls directly into the classification category. We will seek to split the inbound signals into various subsets based on a variety of different attributes to form a
classification tree, which is also referred to as a decision tree. Specifically, we will use classification trees to classify all instances (inbound arrivals) to a predefined set of arrival behaviors (early arrival, on-time arrival, or late arrival) based on different attribute values (such as ship mode, destination, signal source, etc.).

Classification trees are a popular technique in data mining due to their simplicity and transparency. Rokach and Maimon list many other benefits of using classification trees including:

- Versatility for a wide variety of data mining tasks
- Self-explanatory and easy to follow
- Flexibility in handling a variety of input data: nominal, numeric, and textual
- Adaptability in processing datasets that may have errors or missing values
- High predictive performance for a relatively small computational effort

Classification trees are formed with a series of recursive nodes and edges. The first node or “root” will have no incoming edges, while all other nodes or “leaves” have exactly one incoming edge. Each node corresponds with a specific attribute and the outgoing edges then correspond to a specific attribute value.

![General classification tree structure](image)

**Figure 7. General classification tree structure**

---

14 Rokach, Data Mining with Decision Trees.
15 Rokach, Data Mining with Decision Trees.
Instances are then classified by navigating from the root to the final leaf or terminal node based on the attributes that the instance most closely maps to. Classification trees are then “grown” using a training set so that terminal node in each branch corresponds with a historically observed result.

2.3 Attribute Selection

More important than the general hierarchical structure of classification trees is the way in which attributes are selected. Guyon and Elisseeff describe the three main objectives of variable selection as:

1. Improving the prediction performance of the predictors;
2. Providing faster and more cost-effective predictors; and
3. Providing a better understanding of the underlying process that generated the data.  

The figure below illustrates the importance of choosing the right attribute:

![Figure 8. A two class example with independently and identically distributed (i.i.d.) variables](image)

In Figure 8, unless one was aware of the two different classes of data, it would be easy to assume that the entire data set was normally distributed over zero. Only by exposing the existence of the two classes can we see that the data set is really two different distributions with very different means.

---

16 Guyon and Elisseeff, “An Introduction to Variable and Feature Selection.”
17 Guyon and Elisseeff, “An Introduction to Variable and Feature Selection.”
In this thesis, we will build our model based on the assumption that subsets of signals with shared characteristics will perform similarly over time. In other words, we expect that certain subsets will have a tendency to arrive early, arrive on-time, or arrive late. For example, a certain vendor may tend to delivery their shipments late or a particular carrier could have a fantastic track record for delivering on time.

Because we are interested in grouping the signals by the similarity of past performance in terms of estimated delivery date accuracy, we will select the attributes that minimizes the overall joint variance of the model's prediction errors. If the subsets of the data set $S$ are indexed $i = 1, \ldots, k$, then the joint variance ($\sigma_{joint}^2$) is calculated by the weighted average of the subset variances $\sigma_i^2$:

$$\sigma_{joint}^2 = \frac{\sum_{i=1}^{k} n_i \sigma_i^2}{\sum_{i=1}^{k} n_i}$$

Equation 1. Joint variance equation
3. The Inbound Process

3.1 An Overview

Amazon’s supply chain can be divided into two segments: the inbound supply chain and the outbound supply chain. The inbound supply chain consists of all the steps necessary to make a product available for a customer to purchase, while the outbound supply chain consists of all the steps necessary to fulfill an order placed by a customer. Below is a high-level diagram of Amazon’s supply chain:

![Diagram of Amazon supply chain]

Internally, the line between inbound and outbound operations is drawn within the Amazon FCs. Inbound operations within the FCs begin with unloading the trucks at the dock and end with the products being stowed into the warehouse shelves, while outbound operations begin with the items being picked off the shelves and end with the items being shipped out from the warehouse to the customer:

![Diagram of FC processes]

Figure 9. Amazon supply chain overview

Figure 10. FC processes
But even before any shipment arrives at an FC, a significant amount of work has been done to plan, procure, and transport the sellable products into the Amazon fulfillment network.

3.2 The Procurement Process

The process of procuring products from vendors to bring into the Amazon network begins with the retail team. The retail teams are organized around different product groups and they manage everything from vendor relationships to maintaining oversight of the in-stock levels on the retail website. The retail teams initiate the inbound supply chain process by collecting item information, vendor information, and generating demand forecasts and sales plans.

All of this information is then fed directly into Amazon’s automated planning system which then uses this information to answer all of the relevant questions necessary to create a purchase order:

![Planning logic and key decisions](image)

Each of the questions posed above represents a functional system that mines both current and past data to determine the most accurate answer. Amazon’s planning system then passes on this information to its own procurement system which sends out the purchase order back to the appropriate vendor over EDI or through Amazon’s vendor website portal:
Amazon has leveraged the use of technology and algorithms to create a robust and automated inventory planning and procurement infrastructure. By maximizing the use of automation, Amazon has made it possible for its retail team to focus on the tasks where automation makes less sense (contract negotiation, vendor-relationship management, problem solving, etc.). The high degree of automation has several key benefits which include:

- **Low Cost** – Automating many of the inventory and procurement decision translates to fewer retail managers that Amazon will have to hire;
- **High Scalability** – Amazon is able to keep up with high year-over-year growth and complexity as both website volume and SKU selection grows since most of the decisions are automated;
- **Systematic Optimization** – Such an automated approach also allows Amazon to optimize the amount and location of its inventory at the item level;
- **Increased Speed and Efficiency** – The high level of automation also translates into a fluid decision making process so that purchases are made in a continuous basis optimized around when Amazon wants the item to arrive into its network.
3.3 The Inbound Transportation Process

When purchase orders are made, Amazon communicates its ship and delivery expectations to vendors by providing either a delivery-window (a date range that vendors must deliver the ordered items by) or a ship-window (a date range that the vendor must ship the ordered items by) depending on who is responsible for coordinating and paying for the transportation of the shipment to the FCs. Prepay freight terms refer to the freight that vendors are responsible for coordinating and paying for the shipment of and must be delivered within the agreed upon delivery-window. Otherwise, vendors will face chargebacks. Collect freight terms refer to any freight that Amazon pays for and coordinates shipment to an FC through its preferred carriers.

For all collect freight term shipments, vendors are required to submit a routing request a day prior to when the shipment is ready to be picked up. All shipments are tendered automatically by Amazon to a list of preferred carriers. Then the carrier is responsible for coordinating with the vendor to pick up the shipment.

![Routing request process diagram](image)

**Figure 13. Routing request process**

Typical third party carriers can be categorized into four transportation modes:

1. **Intermodal (IM)** – refers to shipments sent via rail
2. **Small Parcel (SP)** – refers to shipments sent via small package carriers that handle individual boxes (i.e. Fed-Ex, UPS, USPS)
3. **Less-than-Truckload (LTL)** – refers to palletized shipments not large enough to fill one truck
4. **Full Truckload (FTL)** – refers to palletized shipments that are large enough to fill one truck
The inbound transportation process for all prepaid freight term shipments are very similar to the collect freight term process except that vendors will submit their routing request directly to the carrier of their choice. While this reduces Amazon’s coordinating responsibilities, it also decreases the level of visibility and control that Amazon maintains over the inbound process.

3.4 The Inbound FC Process

The inbound FC process begins with the coordinating efforts initiated by the carriers and ends with the items being stowed on the shelves:

![Diagram of Inbound FC Process]

**Figure 14. Inbound FC process**

Each carrier will make a request to the designated FC on when they would like to deliver their shipment and then the dock clerk at each FC schedules an appointment slot as soon as possible. On the scheduled date and time, the carrier will bring its freight to the FC dock where it will be unloaded and scanned in to note that the shipment has arrived.

From the FC dock, the shipments will be sorted based on size and each item will be individually scanned and placed into a tote. At this point, the item has been “received” into the Amazon network meaning that it will be ready for the customer to purchase soon. Once the tote is full it is placed on a conveyor system which takes it to the stow area where each item is placed directly on the warehouse shelves or bins.
4. Data Collection, Analysis, and Model Development

4.1 Data Collection from Potential Signal Sources

From the inbound visibility perspective, shipments begin when the vendor has confirmed the purchase order and end when the shipment is received at the FC. Based on Amazon's inbound processes, we identified five predictive signals in the inbound process:

- **Signal 1. Amazon Estimated Delivery Date** – The day Amazon anticipates receiving the ordered item based on historical vendor performance and internally generated need-by dates.

- **Signal 2. Initial Vendor Estimated Delivery Date** – When vendors confirm the purchase order, they also include an update based on when they think the item will arrive at the Amazon FC.

- **Signal 3. Updated Vendor Estimated Delivery Date** – After the item has shipped, the vendor will send a shipment confirmation notice and include another updated arrival estimate.

- **Signal 4. Carrier Estimated Delivery Date** – After the carrier has picked up the items from the vendors, they will communicate their own estimated delivery date to Amazon.

- **Signal 5. Carrier Requested Delivery Date** – Once the carrier is near the destination FC, they will submit the date that they would like to make their delivery to the FC.

These five predictive signals are illustrated in the order they are received by Amazon below:
In order to assess the accuracy of each signal source, we compared each estimated delivery date to the actual receive date or the date that the item was scanned into the Amazon network. This difference between the predicted and the actual receive date is the prediction error and is measured in days (with negative days indicating an earlier than expected arrival and positive days indicating a later than expected arrival). For this study we limited the data collection to a small percentage of purchase orders made by Amazon.com between January 2013 and June 2013 for North America (20,000 observations from each signal source).\(^\text{18}\)

Breaking down the inbound signals by signal source reveals a wide range of error profiles:

---

\(^{18}\) These orders excluded backorders and non-routine reorders (new SKU introductions, overseas imports, seasonal items, and long lead time SKUs). This initial data set contained 100,000 purchase orders.
Segregating the signals by signal source reveals some stark differences in signal quality (Signal 1 Variance = 21.74; Signal 2 Variance = 58.07; Signal 3 Variance = 17.99; Signal 4 Variance = 13.01; Signal 5 Variance = 5.40) in addition to some very different distribution profiles.

4.2 Attribute Selection for Model Development

Given the wide range of variances associated with each signal source, building a classification tree based on the signal source as the primary attribute makes sense:
However, this classification tree model only represents a slight improvement over not classifying the signals at all when comparing the joint variance of the classification tree model to the variance of prediction errors for the complete data set:

\[
\sigma^2_{\text{classification by signal source}} = \frac{20000(21.74 + 58.07 + 17.99 + 13.01 + 5.40)}{20000 + 20000 + 20000 + 20000 + 20000} = 23.24
\]

\[
\sigma^2_{\text{no classification}} = 24.66
\]

To gain a more accurate understanding of the unique variation profile of each signal, we need to further classify each signal based on other attributes. For example, it could be entirely possible that Amazon makes more accurate predictions for certain product groups or that certain carriers make more accurate delivery predictions than others.

We can identify these ideal classification attributes by finding the attribute that minimizes the joint variance for each signal source. But first we need to determine which attributes are appropriate to include in this analysis. Because the underlying distribution of errors for our predictions (when broken up by signal source) is not normally distributed, we opted to use the Kruskal–Wallis one-way Analysis of Variance (ANOVA) which does not assume a normal distribution of the residuals. The null hypothesis tested states that:

There is no difference in the variance of signal X when classified by attribute Y.

And the null hypothesis is rejected when the resulting p-value is less than 0.05:

---

19 Learning and Understanding the Kruskal-Wallis One-Way Analysis-of-Variance-by-Ranks Test
Table 1. Key attributes identified by signal source

<table>
<thead>
<tr>
<th>Signal</th>
<th>Product Category</th>
<th>Payment Type</th>
<th>Vendor</th>
<th>FC Destination</th>
<th>Ship Mode</th>
<th>Carrier Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
</tr>
<tr>
<td>4</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
</tr>
<tr>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
<td>Reject Null Hypothesis</td>
</tr>
</tbody>
</table>

Table 1. Key attributes identified by signal source

After determining the appropriate attributes to include in our analysis, we calculated the corresponding joint variance for each attribute (with the lowest joint variance attribute highlighted for each signal):

Table 2. Resulting joint variance by signal source and attribute (with the lowest joint variance attribute highlighted for each signal)

<table>
<thead>
<tr>
<th>Signal</th>
<th>Signal Error Variance</th>
<th>Product Category</th>
<th>Payment Type</th>
<th>Vendor</th>
<th>FC Destination</th>
<th>Ship Mode</th>
<th>Carrier Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21.74</td>
<td>20.76</td>
<td>21.63</td>
<td>13.58</td>
<td>19.71</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>58.07</td>
<td>56.52</td>
<td>57.94</td>
<td>24.10</td>
<td>56.80</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>17.99</td>
<td>16.81</td>
<td>17.98</td>
<td>9.72</td>
<td>15.75</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>13.01</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>11.20</td>
<td>10.34</td>
<td>8.64</td>
</tr>
<tr>
<td>5</td>
<td>5.40</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>3.28</td>
<td>4.94</td>
<td>4.65</td>
</tr>
</tbody>
</table>

Table 3. Payment Type attribute and Signal 1 results

<table>
<thead>
<tr>
<th>Payment</th>
<th>Number of Observations</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect</td>
<td>18684</td>
<td>20.97439</td>
</tr>
<tr>
<td>Prepaid</td>
<td>1316</td>
<td>30.92123</td>
</tr>
</tbody>
</table>

Table 3. Payment Type attribute and Signal 1 results

And then found the joint variance for these two subsets:
$\sigma^2_{Signal\ 1\ by\ Payment\ Type} = \frac{18684 \times 20.97 + 1316 \times 30.92}{18684 + 1316} = 21.63$ 

Extending our classification tree based on the attributes selected from the above analysis further reduces the joint variance of the model to 11.86—less than half of the variance of the original data set. This represents the final model we will be using to expose the underlying distribution of the errors associated with the inbound signals:

![Classification Tree](image)

**Figure 18. Classification by signal source and lowest joint variance attribute**

With this classification tree, we will subset all of the inbound signal prediction errors from our sample by signal source then by vendor code, carrier code, for FC destination. The end result is a vector of prediction errors for each unique combination of attributes found at each of the terminal nodes. Here is a sample distribution from a vendor in Signal 2:

![Sample Terminal Node Distribution](image)

**Figure 19. Sample terminal node distribution**

We can see that this vendor historically delivered most of their items early, meaning that overall this vendor tends to over predict their delivery dates when confirming the shipment orders.
4.4 Distribution Selection

Although we have determined the attributes with the best predictive ability, we still have not determined how to characterize the resulting error distributions. Based on a random sampling of different terminal node distributions, we observed that they typically fall into one of three categories:

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Description</th>
<th>Sample Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>A continuous probability distribution with parameter ( \mu ) (mean) and ( \sigma ) (standard deviation).</td>
<td><img src="image1" alt="Histogram" /></td>
</tr>
<tr>
<td>Gamma</td>
<td>A continuous probability distribution with shape parameter ( k ) and a scale parameter ( \theta ).</td>
<td><img src="image2" alt="Histogram" /></td>
</tr>
<tr>
<td>Exponential</td>
<td>A continuous probability distribution with rate parameter ( \lambda ).</td>
<td><img src="image3" alt="Histogram" /></td>
</tr>
</tbody>
</table>

Table 4. Commonly observed terminal node distributions

As a result, no single continuous distribution appears to be a good fit for all cases.

Therefore, we recommend that we eliminate any assumption of distribution type and instead utilize the empirical distribution for all cases. In the empirical distribution, the probability of the prediction errors is tied directly to the relative frequency of those errors. As a result, we will make no further assumptions about its underlying distribution and simply assume that all future signals with similar attributes will follow the same distribution of historically observed errors.
4.5 Using the Classification Tree

Once the classification tree has been trained with the sample data and each node now contains a collection of historical prediction errors, using it to expose the variability of future signals is simply a matter of matching the new signal with the classification tree node that most closely matches its attributes, starting at the top of the classification and working our way down. Once we reach a node where there can no longer find matching attributes below, we have found our closest match.

In some cases, this will result in a failure to match the new signal with a terminal node. For example, when a new vendor is introduced, we would have no historical basis for predicting the behavior of his vendor. In this case, we will simply use the distribution of errors for the entire signal source to characterize variability of the new signal.

The classification tree can also be modified to remove sections of the tree that have too few data points to make any reliable assumptions about the future variability of the signal. This is done though a process called pruning by preventing the model from splitting the data points if the number of cases in the terminal node is less than some minimum constraint.
5. Classification Tree Model Validation

5.1 Testing Parameters

The primary output of our model is the associated error distribution for each possible inbound signal. Using these distributions, we can now generate new estimated delivery dates that reflect the desired confidence level. Essentially, we are answering the question, “For each inbound shipment, what is the new estimated delivery date with an arrival-by probability of X?” For example, if a vendor provides an estimated delivery date (Signal 2) of 2/1/2014 with the following associated error distribution:

![Graph of error distribution](image)

Figure 20. Sample error distribution (Signal 2)

Then we can not only generate distribution of likely arrivals, but we can also pinpoint a new estimated delivery date with a desired probability that the shipment will actually arrive by this date (in this example the shipment will arrive by 2/2/14 with a 90% likelihood):

![Graph of arrival distribution with shaded area](image)

Figure 21. Sample arrival distribution with the shaded area representing 90% of the distribution
Therefore, our primary metric for measuring the accuracy of our proposed classification tree model is to measure the error between the percentage of shipments the model predicts will arrive by the new estimated delivery dates and the actual arrival dates within the simulation for a wide range of arrival probabilities. For example, if we generate new estimated delivery dates with a 95% likelihood of arrival, then we can reasonably expect that 95% of our shipments will arrive by these dates. Since our proposed classification tree model was trained on sample inbound data from January to June 2013, we used shipment data from orders that arrived in July 2013 to test its efficacy.

5.2 Simulation Assumptions and Limitations

To conserve the confidentiality of Amazon’s shipment data, only a small percentage of data was taken to conduct the simulation outlined in this thesis. As a result, the primary limitation of this study is the lack of observations in many of the terminal nodes for our classification tree (Only ~30% of the terminal nodes had more than 32 observations). However, we opted not to prune the original model because we would not have done so had this model been created with a complete data set. Due to the small number of observations in many of our terminal nodes, we instead made a simplifying assumption that all distributions were normal.

Another key assumption made in this simulation is that early arrivals are not as detrimental as late arrivals. While this is certainly not true for all cases, late arrivals do have the potential to negatively impact customer satisfaction since they increase the probability of stocking out. This is why we focused on the “arrival-by” date (which allows for early arrivals) as opposed to the “arrive-on” date (in which neither early nor late arrivals are acceptable) for our primary metric.

There are also a few special cases in which the proposed model will not work very well:

Special Case 1: Seasonality – Many product groups like books and toys are seasonal items. As a result, the lead time patterns associated with each of these products can differ significantly based on the time of year they are ordered. Because this model is trained on the first half of the year, no element of seasonality is effectively captured.
Special Case 2: Non-Arrivals – Some orders placed by Amazon and confirmed by the vendors may never arrive. This could be the result of a processing error or a miscalculation by a vendor who confirmed items that were out-of-stock. No logic exists to deal with shipments that were expected to arrive a long time ago, but still has not arrived. In reality, depending on how long the shipment has not arrived, the probability of it never arriving also increases. Further analysis is necessary to understand where the cutoff should occur and how this a cancelled PO will affect future orders.

Special Case 3: Long Term Buys – Many purchases are planned and sent to the vendors well in advance of the day they are to be shipped to an FC. This may occur with new product releases and items that are associated with specific seasons like clothes or holiday products. Because the vendors confirm these orders with estimated delivery dates that can be months to a year away from the order date, the error distribution associated with these types of orders may be significantly different from orders that are fulfilled by vendors immediately upon receipt of the POs. Correctly identifying and modeling these types of orders could increase the accuracy of the model overall.

5.3 Simulation Results

We measured the accuracy of our proposed classification tree model by running multiple simulations with each simulation generating new arrival predictions based on a different expected arrival probability. These expected arrival probabilities ranged from 50% to 98%. In addition, we compared these results to the base model that reflects the error distribution of all the signals aggregated together without any classification. The classification model had an average error rate of 3.2% compared to 5.5% of the base model:
For example, in the simulation with the selected confidence level of 1, we expect that both the base model and the classification model to generate new "arrive-by" dates with an 84.1% likelihood of arrival. From Figure 22, we see that for the dates generated by the classification model in the simulation, 86.5% of the shipments arrived by these dates and for the dates generated by the base model, 89.1% of the shipments arrived by these dates.

In addition, it appears that the classification tree models tend to over predict the percentage of shipments likely to arrive for expected arrival probabilities between 50 and 87%, but under-predicts the percentage of shipments likely to arrive for expected arrival probabilities over 87%. This is due to the wide spread of errors associated with many of the signals (See Figure 16. Boxplot of prediction error by signal source) and the normality assumption we made for this model.

Another way to assess the effectiveness of these predictions is to measure the absolute error between the "arrive-by" date generated by the model and the actual shipment receive date. As evidenced below, the classification tree model has on average 10.5% less error when compared to the base model.
However, it is important to note that higher confidence levels increase the average error in an exponential manner.

5.4 Using the Proposed Classification Tree Model

Because the user sets the expect arrival probability, this model can be tailored to a wide variety of uses. Some business users may be interested in keeping a selection of SKUs in stock at all times and can use this model to identify any shipments that have a high risk for arriving late based on current signals and historical patterns. The difference between the original estimated delivery date and the new estimated arrival date can be used to quickly assess the reliability or expected variance of the latest inbound signal since we expect that the delivery date from the model would not differ greatly from the original delivery date if the signal has been historically accurate.

In our proposed classification tree model, shipments that do not receive any updates are not treated any differently from those that are regularly updated. However, in reality, shipments that do not receive updated signals should be of concern as the lack of a signal may indicate that something has interrupted the inbound shipment process or the flow of information. We can even take this a step further and use this opportunity to proactively reach out to the appropriate party to determine if we need to revise
our estimates. For example, if vendor X needs at least three days of transit time to ship an item to the designated FC, but has not confirmed shipment of the order two days before the EDD, then we can flag this shipment as having a higher risk of arriving late to the appropriate product managers.

5.5 Simulation Conclusions

Overall, the simulation results are promising, particularly since it represents just a small fraction of the inbound data available to Amazon. If these simulations were done again with six months of complete data, we would expect the results to improve even further as a more accurate arrival distribution could be built at every potential terminal node and the recommended empirical distribution is used instead of a normal distribution. As the accuracy of this model improves, we could also potentially answer the question, “On what specific date do we expect this shipment to arrive?” as opposed to “When will this shipment arrive by?”

Yet a broader limitation of this model may hinge on the assumption that past observations is a good basis for future expectations. While the model can be used to differentiate between reliable and unreliable vendors, an improper use of this data could result in perpetuating unwanted behavior. So in the final section, we will address this issue and outline some steps Amazon can take to eliminate some of the variation that we modeled in this thesis.
6. Beyond Tracking Inbound Signals

The model in the previous sections strives to improve prediction on estimated delivery dates for inbound shipments given the current levels of variability in the inbound process. An equally important question is what levers Amazon can utilize to influence supplier behavior and eliminate some of the variability. We will address two such levers in this section: vendor compliance policies and supplier-retailer partnerships.

6.1 The Role of Vendor Compliance

In supplier-retailer relationships, punitive reinforcement is typically manifested in the retailer’s vendor compliance policies. Wal-Mart remains notorious for having one of the strictest vendor policies. One CEO noted that, “Everyone from the forklift driver on up to me, the CEO, knew we had to deliver [to Wal-Mart] on time. Not 10 minutes late. And not 45 minutes early, either. The message came through clearly: You have this 30-second delivery window. Either you’re there, or you’re out.”\(^{20}\) Since Wal-Mart usually commands the lion’s share of a retailer’s sales volume, this threat effectively drives high rates of compliance. More commonly though, a vendor compliance policy generally outlines a schedule of chargebacks (additional costs levied onto suppliers for failing to meet previously agreed upon standards).

For Amazon, the role of vendor compliance is arguably more complicated than it is for Wal-Mart. While Wal-Mart has the economic leverage and substitutability (many suppliers competing for contracts) to be picky with their suppliers, Amazon’s emphasis on selection as part of their strategy means that they must be both willing and lenient enough to work with thousands of vendors with vastly differing levels of capabilities.

Yet Amazon’s continued growth necessitates a second look at their policies. Although we will not discuss the specifics of Amazon’s policies, we will outline some points that should be considered:

1. Should Amazon leverage its growth to demand a stricter adherence to its policies?

2. Should Amazon be willing to cut out poor performing suppliers?

3. Does the current chargeback system capture the true total cost incurred by Amazon?

How Amazon adapts its vendor policies will either exacerbate or alleviate the rate at which Amazon will have to continue to build out its FC network.

6.2 Supplier-Retailer Partnerships

While vendor policies place emphasis on the punitive elements of the supplier-retailer relationship, there is a tremendous opportunity in growing positive long-term strategic partnerships. Many companies in both retail and non-retail industries have gained a distinct competitive advantage through strong cooperative relationships with suppliers. Japanese automakers such as Toyota have long established a successful history based on developing such cooperative relationships characterized by high levels of trust, two-way information sharing, direct assistance by buyers to suppliers to help them improve production performance, long-term supplier contracts, and formal evaluation of supplier performance. 21

For Amazon, one of the main challenges of building such strategic partnerships is the sheer size of its supplier base. Building such relationships takes time and a significant expenditure of both human and monetary resources. But Amazon has already proven itself to be capable of launching partnerships with initiatives such as the Amazon Certified Frustration-Free Packaging. 22 With Frustration-Free Packaging, Amazon worked directly with manufacturers to redesign the packaging material to be easy to open, recyclable, and reduce the overall amount of packaging waste. 23

Our recommendation is to focus on the following:

1. Vendor Segmentation – Given the large supplier base that Amazon must work with, segmenting the vendors based on different profiles allows Amazon to focus its efforts appropriately. One of the benefits of the proposed classification tree model is that it provides a way of measuring vendor performance as a function of delivery reliability. Using historical performance, the large

21 Langfield-Smith and Greenwood, “Developing Co-Operative Buyer–Supplier Relationships.”
22 “Amazon.com Help: About Amazon Certified Frustration-Free Packaging.”
23 “Amazon Frustration-Free Packaging @ Amazon.com.”
vendor base can be segregated to identify those that are historically highly reliable and those that are not. Other elements Amazon should consider are the risk of stock out, average lead time length, and even the age of the relationship.

2. **Scalable Solutions** (like automated feedback loops for suppliers) – By notifying suppliers when they are in danger of missing their target delivery dates, Amazon may be able to prevent some late deliveries while also alerting the supplier that their shipment is being carefully watched.

3. **Vendor Education** – Building awareness and knowledge of how both early and late shipments impact customer fulfillment in addition to a formal channel for providing feedback on past performance may reduce unwanted vendor behavior.

4. **Performance-based Incentives** – In the long run, Amazon may find it valuable to categorize and recognize well-performing vendors through some sort of vendor-certification program. In addition to the recognition, Amazon will need to offer some tangible incentives (pricing, long-term contracts, etc.) to generate interest in its supplier base.

**6.3 The Path Ahead**

When working with thousands of vendors with widely varying capabilities, it may not be feasible to have a “one-size fits all” policy. Instead, Amazon could potentially lower or raise its lead-time expectations to match the vendors’ best inherent capabilities. How Amazon can categorize or separate these vendors should be the subject of another research project.

Additional thought and research should be done on the trade-off between on-time delivery performance and procurement lead time. The bulk of the research conducted in this project was around variation in delivery performance. It is important to note that eliminating the variation in delivery performance does not eliminate the variation in lead time. Regardless of whether the shipment arrived on time, the amount of total lead time will still differ significantly based on vendor locations, vendor turnaround times, and the ship mode.
The relationship between delivery performance and acceptable procurement lead time remains unclear. For example, if Amazon placed great emphasis on delivery performance and significantly increased chargebacks for missing a delivery, it may be possible that we would see lead times rise over time as vendors build in more buffer into their projected lead times to reduce the opportunity for a late delivery. If this turns out to be true, then the next questions is, “Which is worse, longer lead times but higher on-time delivery performance or shorter lead times but lower on-time delivery performance?” More importantly, we need to know if it is possible to drive lead times down without seeing on-time delivery performance deteriorate.

We write this to remind ourselves to not take a myopic view of the metrics we use to drive performance in our supply chains. The pressure we may exert to “fix” one part of the supply chain may have significant unintended consequences elsewhere.
Bibliography


Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. “From Data Mining to Knowledge Discovery in Databases.” *AI Magazine* 17, no. 3 (1996): 37.


