ANALYSIS AND OPTIMAL DESIGN OF DISCRETE ORDER PICKING TECHNOLOGIES ALONG A LINE

Donald D. Eisenstein

Graduate School of Business, The University of Chicago,
Chicago, Illinois 60637 USA.
don.eisenstein@chicagogsb.edu

Draft: December 27, 2006
Work in process: Please do not cite without permission

Abstract

Order picking accounts for most of the operating expense of a typical distribution center, and thus is often considered the most critical function of a supply chain. In *discrete* order picking a single worker walks to pick all the items necessary to fulfill a single customer order. Discrete order picking is common not only because of its simplicity and reliability, but also because of its ability to pick orders quickly upon receipt, and thus is commonly used by e-commerce operations. There are two primary ways to reduce the cost (walking distance required) of the order picking system. First is through the use of *technology* — conveyor systems and/or the ability to transmit order information to pickers via mobile units. Second is through the *design* — where best to locate depots (where workers receive pick lists and deposit completed orders) and how best to lay out the product. We build a stochastic model to compare three configurations (two in common use, and a third a new proposed configuration). And for each configuration we explore the optimal design.

Within the warehouse, the order picking function typically accounts for about 55% of operating costs (see Frazelle, 1996). In general, order picking is commonly considered the most critical function in a supply chain (see for instance Tompkins and Smith, 1998). The importance of order picking is becoming more apparent as new e-commerce operations struggle to compete with traditional bricks-and-mortar operations. Consider for example e-commerce grocery services (such as Webvan or Peapod). They distinguish themselves from traditional grocery stores in that they must absorb the cost to pick customer orders, whereas for a traditional grocery store the customer performs this function for free.

In *discrete* order picking a single worker picks all items necessary to fulfill a single customer order, and picks no other items until the order for the customer is complete. This method of order picking is common because it is simple and reliable — a picker need manage only one customer order at a time. Furthermore a customer order is picked quickly upon receipt without delaying to batch with other customer orders, or delaying to hand off a partially picked order from one picker to another; and therefore, discrete order picking is commonly used for real-time operations. For instance discrete order picking is used at some discount stores such as Service Merchandise, where essentially all items for a customer order are picked from the in-store warehouse while the customer waits.

Figure 1 depicts an implementation of discrete order picking along a linear pick line. The picker retrieves a printed pick list for a single order at the *depot*, located at the beginning of the line, along with an empty bin to carry the items. He walks as far down the line as necessary to retrieve all the items ordered by the customer, and then returns to the depot to deposit the picked items and retrieve a new pick list and empty bin. Other workers may also be simultaneously and independently picking orders.

One implementation of discrete order picking along a line, in which the author has been involved, is at Urbanfetch.com, an e-commerce retailer that delivers a large variety of items to customers within an hour. Discrete order picking is the natural picking discipline for Urbanfetch. Once an order is received it must be picked quickly. A worker retrieves the pick list at the depot, walks to retrieve all items, and then returns the items to the depot where they await dispatch for delivery. Perhaps most importantly, discrete order picking assures high quality, as a single picker is dedicated to just one order at a time. Management is also able to easily adjust picking capacity on the fly, moving workers to and from restocking or other areas when needed (even senior management picks orders when necessary). Thus the advantages for Urbanfetch is that the picking protocol is simple, reliable, scalable, and responsive.

Since it is advantageous for workers to avoid the waste associated with walking between aisles or across conveyor systems, discrete order picking along a line is common. Furthermore, discrete order picking along a line acts as a fundamental building block of other more elaborate picking schemes. Consider for example a *zone* picking system as shown in Figure 2. Here, five workers are each assigned to their own zones (for

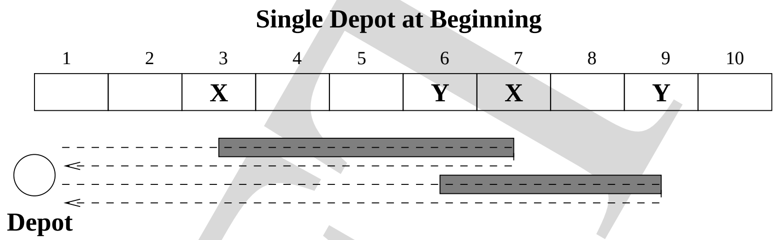

Figure 1: Two customer orders, X and Y. Order X has its leftmost pick at location 3, and its rightmost pick at 7. Order Y has its leftmost pick at 6, and its rightmost pick at 9. With a single depot located at the beginning of the line the picker walks along the line until the order is picked. The shaded portion depicts the part of the walk path from when the first item for the order is picked until the last.

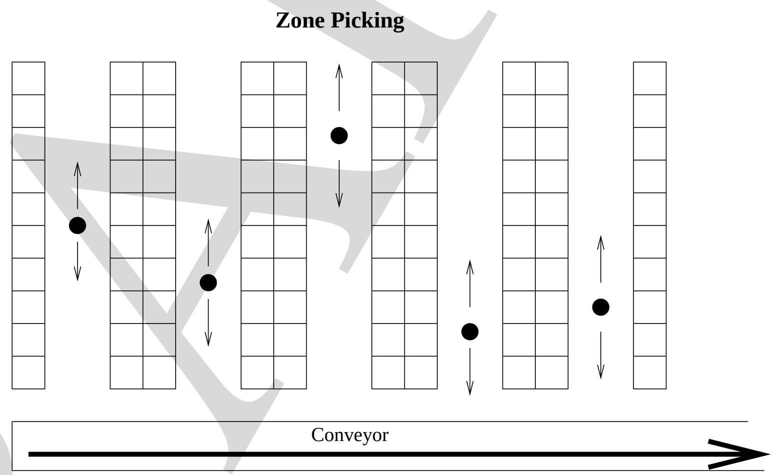

Figure 2: Each of five workers remain in their zones (in this case an aisle). Here, this zone picking operation is composed of five, independent, discrete picking operations.

simplicity each zone is shown as its own entire pick aisle, but it is common for a worker zone to cover only a portion of an aisle). Each picker remains in his zone, picking just the items that are necessary for each customer order. When his picks are complete within his zone he deposits the items on the conveyor. Each completed order is then carried by conveyor to perhaps a consolidation area before eventually arriving at the shipping department. (See Tompkins and Smith, 1998, for an overview of order picking protocols including discrete and zone picking; and Bartholdi et al., 2000, for an implementation where each worker picks a portion of a pick line.)

The main disadvantage of discrete order picking is that the amount of walking per pick can be high. Multiple customer orders can however be combined into a batch (*batch* picking) to help reduce the walking per pick, and our models equally apply to such batch picking operations. Our focus here is to examine how to configure the system to minimize the amount of walking between picks. We do not consider the time to actually pick an item once the picker is in position, since this time is required (fixed) regardless of the configuration. But we note that technologies such as pick-to-light systems can speed up this fixed portion of the efforts of the picker equally well for all of our configurations.

We consider two primary ways in which one can reduce the walking required for a discrete order picking operation.

Configuration: The first is through the *configuration* of the system. By the configuration we mean the way in which the picker receives picking information (the pick list) and the way in which he deposits the items after they are picked. He may receive the pick list in print form at a depot, or more technology can be used so that the pick information is transmitted to him via a mobile device. The picked items of an order may be deposited by walking them back to a depot location; or with conveyor technology, the completed order can be deposited anywhere along the pick line onto a conveyor system.

Layout: The second basic way to impact the performance of a discrete order picking system is through the *layout* of the items. Since some items are more likely than others to be required in an order, the layout should be matched to the configuration to be most effective.

We examine three basic configurations:

- **Single-depot:** with no technology investment,
- **Dual-depots:** with conveyor technology investment, and
- **No-depot:** with conveyor and RF technology investment.

The first configuration is the simplest and most common — a *single depot at location k.* Figure 1 shows a common case where the depot is located at the start of the line, $k = 1$. Figure 3 shows a case when a single depot is located within the pick line. The picker begins each order by walking from the depot to the leftmost pick, and then picking all the items in sequence until the rightmost item is picked, and then walking back to the depot. It is simple to convince oneself that it is never better to alter such a sequence of picks by picking an item central to the order first.

Our next configuration introduces a new design to discrete order picking — locating two depots along the pick aisle (see Figure 4). We have never seen this in practice, except when the two depots are located at both ends of the pick line. The protocol is that a worker would first pick up a pick list at the left depot, and pick the order with a left-to-right pass along the line. He would then visit the right depot to deposit the current order and retrieve a pick list for a new order, and then make a right-to-left pass along the aisle to pick the next order. And then finally return to the left depot to deposit the current order and retrieve a new pick list.

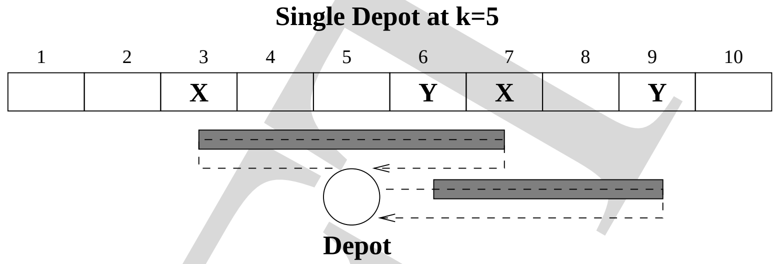

Figure 3: Two customer orders, X and Y. With a single depot located at $k = 5$ the picker first walks to the leftmost pick of order X, then along the line picking order X until the rightmost pick for order X is complete. Then he returns to the depot and repeats the procedure for order Y.

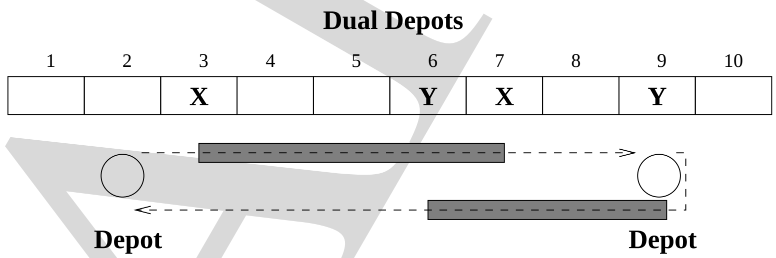

Figure 4: Two customer orders, X and Y. With dual-depots the picker first starts at the left depot and walks to the leftmost pick of order X, then walks along the line picking order X until rightmost pick is completed. He then returns to the right depot, and then walks to the rightmost pick of order Y, then walks along the line picking order Y until the leftmost pick is complete for order Y, before returning to the left depot for the next order.

A centralized computer will simply alternate printing pick lists between the two depots, and when printing a pick list to the leftmost (rightmost) depot it prints the items in a left-right (right-left) sequence so picking is easy. For many implementations this configuration might require a conveyor to transport picked orders from the two depots to shipping. We refer to this configuration as *dual-depots*.

Finally we consider the best possible discrete picking implementation that requires full technology, eliminating the need to walk to and from depot locations. Here, a conveyor is installed so that a picker can deposit a completed order anywhere along the pick aisle. In addition, the picker has a mobile unit that transmits pick information. This is commonly accomplished using an RF (Radio Frequency) device, which is often strapped like a large watch on the wrist. Thus when a worker completes an order, he immediately, wherever he is located, deposits the completed order on the conveyor, then collects an empty bin (which are stacked under the length of the conveyor or along an overhead rail), and then he pushes a button on his RF device to display the picks necessary for a new customer order (see Figure 5). We call this configuration *no-depot*.

We build a stochastic model of the discrete order picking operation. For each configuration we determine the expected amount of walk distance required per order. We do not model the time required to stop and

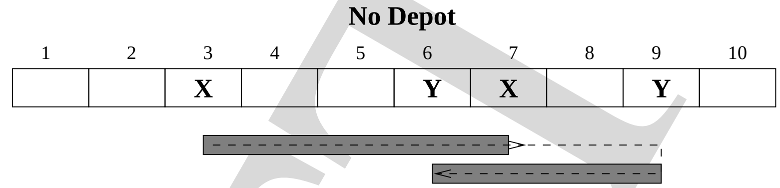

Figure 5: Two customer orders, X and Y. With no depot the picker starts at the leftmost pick of order X, then walks along the line picking order X until the rightmost pick is complete, and then deposits order X along the conveyor. He then walks to the rightmost pick of order Y, then walks along the line picking order Y until the leftmost pick is complete, and then deposits order Y on the conveyor.

pick each item, since this is the same regardless of the configuration — the configurations differ only in the distance walked per order. Our model is useful in ascertaining the benefit of technology — an extra depot, a conveyor, or RF devices. We will find that the benefit depends on the characteristics of the order stream — many pick locations required per order, or very few?

Next we consider design issues for the single and dual-depot configurations. First we will consider where best to locate single or dual depots given a product layout. This is a practical design issue when a depot is simply a printer to output pick lists and a rack to hold picked orders and thus can easily be moved. In contrast, changing the product layout is costly and disruptive, and so is done infrequently.

On the other hand in some facilities, the depot locations are static (for instance they might correspond to shipping docks or the front door of an e-commerce distribution center). In such cases we investigate how best to layout the product when the depot location(s) are fixed. We also examine how to layout items for the no-depot configuration. Finally, we examine the full design question for single and dual depots — how best to both locate depots and layout product simultaneously.

1 Our model

We model a stochastic stream of orders in which the makeup of each order is independently and identically distributed. We divide the pick line into n locations or regions with each order represented as a binary n-tuple, the i-th entry of which represents whether at least one pick is required at location i. The probability of at least one pick occurring at location i is p_i, and is independent of other locations. For instance, if the pick line is composed of twenty static shelves, we might model an order by setting $n = 20$. Then for each order, the probability of at least one pick required from shelf i is p_i, and the probability shelf i requires no pick is $1 - p_i$. We then, for each configuration, determine the expected walk distance required for each order.

Our model does not explicitly handle correlation among the items. However, it is common in practice for strongly correlated items to be located together. For example at Revco Drugs, products are stored for

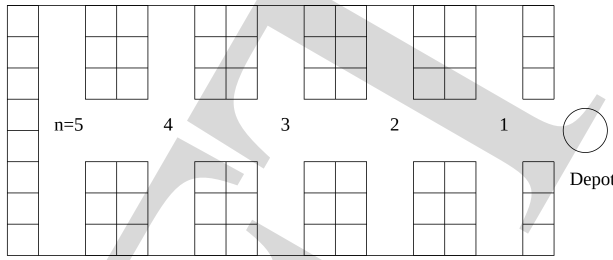

Figure 6: Here we show a common picking layout in which discrete pickers navigate $n = 5$ aisles.

picking in the distribution center in the same way as retail stores are laid out in order to facilitate restocking at the retail stores (see Bartholdi et al., 2000). Therefore, for example, while there is certainly considerable correlation between different shampoos they are located in the same rack, and therefore are represented by the same p_i. Similar items are also often shelved together because of their physical attributes — for example at Urbanfetch.com snack chips require different shelving than video tapes which require still different storage than ice cream. So, in many practical settings, we can expect our model to handle some types of correlation well.

We assume that the distance between adjacent pick locations is the same and of unit length. This simplifies our notation with little sacrifice to the robustness of our model — if different distances between locations is needed then dummy locations can be inserted with p_i's set to 0. A main element of our model is that if a picker travels from location i to location j, then he walks a distance of $j - i$ units. Nevertheless, our model might be used to approximate configurations such as shown in Figure 6 where small aisles cross a single central aisle. Here we might model $n = 5$ and have p_i be the probability a picker will have to visit cross-aisle i (to the left and/or to the right of the main aisle). Thus in this case we model the distance a picker must walk along the main aisle, but not how far into any cross-aisle.

2 Related work

The closest work in the literature appears to be Jarvis and McDowell (1991). They concentrate on the product layout design problem for a single-depot configuration for a discrete order picking system in which a single picker navigates parallel aisles (a rectangular warehouse) like that shown in Figure 2). Since they assume that once a picker enters a vertical aisle he must traverse the entire aisle, the heart of their model is the same as our single-depot model — how far must the picker walk along the horizontal line. They determine the best product layout for two special cases of a single-depot configuration — a depot at the beginning of the line or one in the center. Our work for the single-depot model generalizes their work by considering a single depot anywhere along the pick line.

Chew and Tang (1999) also examine a discrete order picking system in a rectangular warehouse. Their focus is on examining the effects of batch sizes in a real time system. Thus their probability model fixes the number of items per order (the batch size), and then develops a queueing model where incoming orders queue and are batched before picking. Rosenwein (1995) examines heuristics for batching orders.

Bartholdi et al. (2000) examine bucket brigade order picking in which a team of workers combine efforts to pick customer orders.

Another stream of literature considers how best to traverse a rectangular warehouse to pick a given order; Ratliff and Rosenthal (1983) solve this special case of the traveling salesman problem, and Hall (1993) considers the expected route lengths in a rectangular warehouse for a variety of layouts and strategies.

De Koster et al. (2006) provide a nice overview of the order picking literature. Matson and White (1982) provide a general overview of research in material handling, and van den Berg (1999) provides a nice literature review of the planning and control of warehousing systems.

3 Stochastic analysis

We now derive expressions for the expected walk distance required per order for each of our three configurations. The orders are identically and independently distributed, and for a given order we let p_i be the probability that at least one pick is required at location i, and we let $q_i = 1 - p_i$. We let the random variable

$N =$ the total number of locations requiring picks in an order.

Our method of modeling random orders will allow a null order, $N = 0$, to occur, however we will condition all of our expected walking distances on $N > 0$, non-null orders. We denote the probability of a non-null order by

$$\mathcal{P} = P(N > 0) = 1 - \prod_{i=1}^{n} q_i. \tag{3.1}$$

3.1 Single depot at k

We now determine the expected walk distance per order when a single depot is located along the pick line at real valued location k (see Figure 3). The walking required per order has two components: The distance from the depot to the leftmost pick and back to the depot, plus the distance to the rightmost pick and back to the depot. We define two random variables:

$L_k =$ the distance from k to the leftmost pick, and

$R_k =$ the distance from k to the rightmost pick.

Now letting the random variable

$S_k =$ the total walk length per order for a single depot located at k,

we have

$$S_k = 2(L_k + R_k). \tag{3.2}$$

We now define variables that count the number of locations not requiring picks at the beginning of the line and at the end of the line.

$L =$ number of consecutive unvisited locations at the beginning of the pick line,

and

$R =$ number of consecutive unvisited locations at the end of the pick line.

Therefore the unconditional probability mass function for L is given by

$$P(L = i) = p_i \prod_{j=1}^{i-1} q_j \quad \text{for } i = 0, \dots, n,$$

and similarly for R we have

$$P(R = i) = p_i \prod_{j=i+1}^{n} q_j \quad \text{for } i = 0, \dots, n.$$

So we have

$$E[L_k] = \sum_{i=1}^{\lfloor k \rfloor} (k - i) P(L = i), \tag{3.3}$$

and,

$$E[R_k] = \sum_{i=\lceil k \rceil}^{n} (i - k) P(R = i). \tag{3.4}$$

$$\tag{3.5}$$

Now conditioning on non-null orders, $N > 0$, we have:

$$\begin{aligned} E[S_k | N > 0] &= \frac{E[S_k]}{\mathcal{P}} \\ &= \frac{2\left(E[L_k] + E[R_k]\right)}{\mathcal{P}} \\ &= \frac{2}{\mathcal{P}} \left[\sum_{i=1}^{\lfloor k \rfloor} (k - i) P(L = i) + \sum_{i=\lceil k \rceil}^{n} (i - k) P(R = i) \right]. \end{aligned} \tag{3.6}$$

3.2 Dual depots

We now determine the expected walk distance per order with two depots, one located at u and the other at v, where $u \leq v$. With two depots we assume that the picker alternates between the depots (see Figure 4). Although one can create instances where the picker would walk less if he deviated from our protocol by visiting the same depot on consecutive orders, we assume a simple and easily implementable protocol that alternates orders between the two depots.

We can consider the walking required per order to have three components: The distance from the left depot to the leftmost pick and back to the left depot, the distance from the left depot to the right depot, and the distance from the right depot to the rightmost pick and back to the right depot. Then it repeats in reverse. We let the random variable

$D_{u,v} =$ the total walk distance for dual depots at locations u and v.

So we have that

$$D_{u,v} = 2L_u + (v - u) + 2R_v.$$

So the expected walking distance per order for dual depots located at u and v along the pick line is given by

$$E[D_{u,v}|N > 0] = \frac{1}{\mathcal{P}} \left(2E[L_u] + (v - u) + 2E[R_v]\right). \tag{3.7}$$

We note that this expectation is a generalization of the single-depot case. That is

$$E[D_{k,k}|N > 0] = E[S_k|N > 0].$$

3.3 No depot

We now determine the expected walk distance per order with no depot. With no depot we assume that the picker alternates the direction he picks — picking one order left-to-right and the next right-to-left (see Figure 5).

Thus the walking distance required per order has two basic components. First is the distance from the leftmost pick to the rightmost pick (the length of a pick run). Second, there is the distance from the end of one order to the start of the next: from the rightmost pick of one order to the rightmost pick of the next, or between leftmost picks of consecutive orders.

We let

$T =$ the total walk distance with no depot,

and let

$W =$ the length of the pick run.

We define the random variables for the distance from the end of one order to the beginning of the next:

$B_R =$ the distance between the rightmost pick of one order to the rightmost pick of another,

and

$B_L =$ the distance between the leftmost pick of one order to the leftmost pick of another.

So we have

$$T = W + \frac{1}{2}\Big(B_L + B_R\Big).$$

The expected length of a pick run is given by

$$E[W] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Big(|j-i|\; P(L=i)\, P(R=j)\Big),$$

and conditioning out null orders gives

$$E[W|N>0] = \frac{E[W]}{\mathcal{P}}.$$

For the expected distance between left pick-ends we have

$$E[B_L] = \sum_{i=1}^{n} \sum_{j=1}^{n} |i-j|\; P(L=i)P(L=j),$$

and between right pick-ends we have,

$$E[B_R] = \sum_{i=1}^{n} \sum_{j=1}^{n} |i-j|\; P(R=i)P(R=j).$$

To condition out null orders from $E[B_L]$ and $E[B_R]$ we consider two independent orders with N_1 and N_2 picks. So

$$P(N_1 > 0 \cap N_2 > 0) = \mathcal{P}^2.$$

Thus we have

$$E[B_L|N_1, N_2 > 0] = \frac{E[B_L]}{\mathcal{P}^2}$$

and,

$$E[B_R|N_1, N_2 > 0] = \frac{E[B_R]}{\mathcal{P}^2}.$$

And so finally the expected walk distance per order with no depot is given by

$$E[T] = E[W|N>0] + \frac{1}{2}\Big(E[B_L|N_1, N_2 > 0] + E[B_R|N_1, N_2 > 0]\Big). \tag{3.8}$$

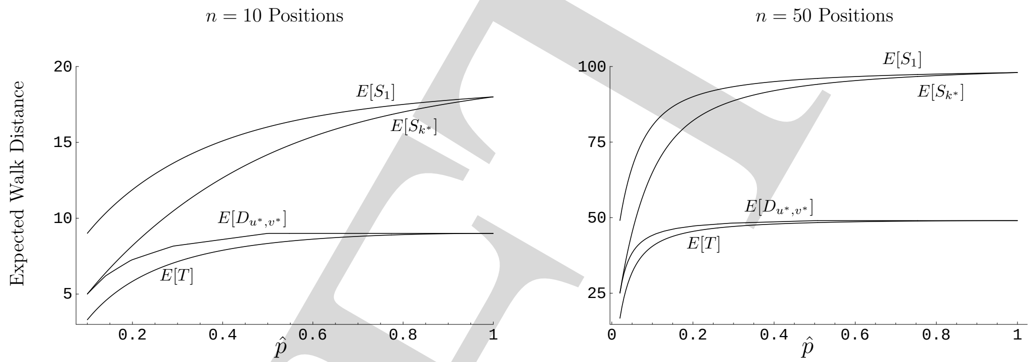

Figure 7: Here we show the expected walk distances for each configuration for both $n = 10$ and $n = 50$. Each item has the same the $\hat{p}$ which we range from $1/n$ to 1.

3.4 Comparing the configurations

Figure 7 shows the expected walk distance with $n = 10$ and $n = 50$ locations for each of our configurations (with the common configuration of a single depot at $k = 1$ included). We denote $\hat{p}_i$ as the probability of a pick at position i after conditioning on non-null orders. For each data point we set each $\hat{p}_i$ to the same $\hat{p}$ value, and this value is then varied along the x-axis from $1/n$ to 1. The $E[S_{k^*}]$ and $E[D_{u^*,v^*}]$ curves are determined using the best depot locations for each $\hat{p}$. As the pick density, $\hat{p}$, increases the amount of walking increases for each configuration. There is a clear dominance relation between the configurations, as indicated by our plots, that holds in general. (We note that we always condition our expected walking distances on $N > 0$, but omit the notation in this section.)

Theorem 3.1. *For any fixed layout and set of p_i's we have $E[S_1] \geq E[S_{k^*}] \geq E[D_{u^*,v^*}] \geq E[T]$.*

Proof. $E[S_1] \geq E[S_{k^*}]$ holds since S_{k^*} is not constrained by the depot location. $E[S_{k^*}] \geq E[D_{u^*,v^*}]$ holds, since, D_{u^*,v^*} is free to locate both depots at k^*. Also $E[D_{u^*,v^*}] \geq E[T]$ holds since the later need not visit an intermediate point between order pick ends. □

The figures also indicate clear limiting behavior. We state the following without proof.

Lemma 3.2. *For any fixed number of locations n, as the pick density increases we have the following limiting* *behavior:*

$$\lim_{\hat{p}\to 1} E[S_1] = 2(n-1)$$

$$\lim_{\hat{p}\to 1} E[S_{k^*}] = 2(n-1)$$

$$\lim_{\hat{p}\to 1} E[D_{u^*,v^*}] = n-1$$

$$\lim_{\hat{p}\to 1} E[T] = n-1.$$

As the pick density increases, $\hat{p} \to 1$, each order tends to require a visit to all locations. In this case a single depot anywhere along the line behaves the same, and poorly, tending to force the picker to walk twice the length of the line for each order. On the other hand, a dual-depot configuration will tend to locate one depot at each end of the line, so that the walking required per order will be tend to be only one length of the line. And similarly, the no-depot configuration will tend to require the picker to walk one length of the line. So when pick densities are large a dual-depot configuration is recommended, performing very well without the added expense of RF technology. We emphasize that two types of behavior emerge — any single-depot configuration requires twice the walk distance of either a dual or no-depot configuration.

We have the following limiting behavior as the pick density gets small.

Lemma 3.3. *For any fixed number of locations n, as the pick density decreases we have the following limiting behavior:*

$$\lim_{\hat{p}\to 1/n} E[S_1] = n-1$$

$$\lim_{\hat{p}\to 1/n} E[S_{k^*}] = \frac{n^2-1}{2n}$$

$$\lim_{\hat{p}\to 1/n} E[D_{u^*,v^*}] = \frac{n^2-1}{2n}$$

$$\lim_{\hat{p}\to 1/n} E[T] = \frac{n^2-1}{3n}.$$

As the pick density decreases, a single depot at the beginning performs very poorly, as pickers must, on average, walk to the center of the aisle and back. But when a single depot is located in the center, the picker must only walk, roughly speaking, one-fourth of the line and back, on average. As the pick density gets very small, optimal dual depot locations will tend toward each other in the center of the line, and thus will tend to behave the same as a single optimal depot. (This is because our dual-depot protocol forces the picker to alternate between the two depots. When the pick density is very small, dual depots would benefit by a more powerful and flexible algorithm.) Finally, the full technology of a no-depot configuration gains a distinct advantage over all others, as it is able to better absorb the uncertainties of fetching small orders.

So as the pick density tends to zero three types of behavior emerge: a single depot at the beginning requires roughly twice the walking required by optimal single or optimal dual depots, and optimal single or optimal dual depots require 1.5 times the walking of a no-depot configuration.

So in conclusion the technology investment should be a function of the characteristics of the order stream. When pick densities are very large, so orders tend to require visits to most the pick locations, then it is important to avoid single-depot configurations. The technology of a conveyor system to facilitate a dual-depot configuration may be warranted. However, the added expense of RF technology for the no-depot configuration is not warranted, it performs no better than a dual-depot configuration.

When orders tend to have very few picks it is important, if a single depot is used, to locate it optimally within the pick line. However, the added technology of using dual depots is unwarranted if the pick densities are very small, but investing in both conveyor and RF technology for the no-depot configuration may indeed be warranted.

But overall, a dual-depot configuration performs very well, yielding significantly to the no depot case only when the pick density is very small. We conclude that in general, the addition of conveyor technology is more critical than the addition of RF technology to reduce picking costs.

4 Optimal design

We will first consider where best to locate single or dual depots given the item layout. Then we consider how best to layout the items given the location of the depot(s), or with a no-depot configuration. Finally, we consider the total design problem of how best to simultaneously locate the depot(s) and layout the items.

4.1 Optimal single and dual-depot locations

We first consider the single-depot case. The next theorem shows how to determine the optimal depot location, k^*, that will minimize $E[S_k|N > 0]$, given by Equation (3.6). We find that k^* is a *median* location. In our case this means that k^* is integral, corresponding exactly to a pick location. And furthermore that, roughly speaking, we choose k^* so that the probability that the picker must walk to the left of the depot is equal to the probability that the picker must walk to the right.

Theorem 4.1. *The expected walk distance per order for a single depot is minimized when the depot is located at*

$$k^* = \min\left\{k \;\middle|\; \sum_{i=1}^{k} P(L = i|N > 0) - \sum_{i=k+1}^{n} P(R = i|N > 0) \geq 0\right\} \tag{4.1}$$

Proof. This is a simple extension to the rectilinear facility location problem on a line (see Francis et al. 1992). Here we seek to locate a new "facility" of minimal weighted distance to n existing "facilities", where existing facility i is located on the line at point i. The weight for existing facility i is $P(L = i)$ if the new facility is located to the right of i and of weight $P(R = i)$ if the new facility is located to the left. □

We now consider the dual-depot case. We seek to find the optimal depot locations u^* and v^* that will minimize $E[D_{u,v}]$, given by Equation (3.7). As in the single-depot case, we again find that both u^* and v^* are median locations. Furthermore we can find the location u^* that minimizes $E[L_u]$, and independently the location v^* that minimizes $E[R_v]$.

Theorem 4.2. *The expected walk distance per order for dual depots is minimized when the left depot is located at*

$$u^* = \min \left\{ u \;\middle|\; \sum_{i=1}^{u} P(L = i|N > 0) \geq \frac{1}{2} \right\}$$

and the right depot is located at

$$v^* = \max \left\{ v \;\middle|\; \sum_{i=v}^{n} P(R = i|N > 0) \geq \frac{1}{2} \right\}$$

Proof. The proof here follows in a similar way as the proof of Theorem 4.1. We find u^* by considering an existing "facility" i to have weight $2P(L = i)$ when the new facility is to the right of i, and we add a dummy existing facility at location n with weight 1 representing the travel between depots. We then find v^* in a similar fashion. □

We note that,

Theorem 4.3. $u^* \leq k^* \leq v^*$.

Proof sketch $\sum_{i=1}^{k} P(L = i|N > 0) + \sum_{i=k+1}^{n} P(R = i|N > 0) \geq 1$ for any k since every order requires either a pick at or to the left of k, or a pick to the right of k — or both. And by the definition of k^* we know $\sum_{i=1}^{k^*} P(L = i|N > 0) \geq \sum_{i=k^*+1}^{n} P(R = i|N > 0)$, and thus $\sum_{i=1}^{k^*} P(L = i|N > 0) \geq 1/2$, and therefore $u^* \leq k^*$. An analogous argument shows $k^* \leq v^*$. □

And following our observations from Lemmas 3.2 and 3.3, as the picks become more dense $u^* \to 0$ and $v^* \to 1$, and as the picks become more sparse $u^* \to k^*$ and $v^* \to k^*$.

4.2 Optimal layout

We are given n entities that we wish to assign to n locations in order to minimize the expected walk distance. We assume that the depot location(s), if present, are fixed. The n entities may be at the item (sku) level, or some larger entity, like a shelf or rack of items, but we will use the terminology that each entity is an item. For each item j we have a probability c_j that no pick is required for item j in a random order. We seek to find a complete assignment that sets each q_i (the probability of no pick required at location i) to a c_j, for $i, j = 1, \ldots, n$, that minimizes the expected walk distance. For convenience we sequence items according to their c_j values so that $c_1 \geq c_2 \geq \cdots \geq c_n$.

4.2.1 Single-depot item layout

Jarvis and McDowell (1991) solve for the best item layout for two special cases — when a single depot is at the start of the line, and when the depot is in the center. Our contribution here is a branch and bound procedure for the general case. In particular we derive a result that characterizes optimal layouts that not only provides useful cuts in our branch and bound algorithm, but solves exactly and quickly the special cases of Jarvis and McDowell.

The conditional expected walking distance per order for a single depot at an integral location k, Equation (3.6), can be rewritten as

$$E[S_k|N > 0] = \frac{2}{\mathcal{P}} \left((n-1) - \sum_{i=1}^{k-1} \prod_{j=1}^{i} q_j - \sum_{i=1}^{n-k} \prod_{j=1}^{i} q_{n+1-j} \right). \tag{4.2}$$

So we have $k-1$ locations to the left of the depot, and $n-k$ to the right. We now establish our first characterization of optimal layouts. If we consider the assignments to the left (or right) of the depot in an optimal layout, then a more popular item must be closer to the depot than a less popular item.

Theorem 4.4. *Any optimal layout with a single depot located at k must assign items so that $q_1 \geq q_2 \geq \cdots \geq q_{k-1} \geq q_k$, and $q_k \leq q_{k+1} \leq \cdots \leq q_{n-1} \leq q_n$.*

Proof. To minimize Equation (4.2) we wish to maximize the terms

$$\sum_{i=1}^{k-1} \prod_{j=1}^{i} q_j + \sum_{i=1}^{n-k} \prod_{j=1}^{i} q_{n+1-j}.$$

Consider the first term of items left of the depot. If one claims to have maximized this expression and yet there is some $q_i < q_j$ with $1 \leq i < j \leq k$, a simple swap will, in contradiction, increase the expression. An analogous argument holds for the second term, for items right of the depot. □

Theorem 4.4 leads immediately to the following special case that was established in Jarvis and McDowell.

Corollary 4.5. *When a single depot is located at the start of the line, $k = 1$, the layout where $q_1 = c_n \leq q_2 = c_{n-1} \leq \cdots \leq q_n = c_1$ will minimize $E[S_1|N > 0]$.*

Theorem 4.4 means that we can view our layout decision problem as a partitioning problem instead of a sequencing problem — once we decide which c_j to assign left of the depot and which to the right, then the sequence is known, they are sequenced with the smaller c_j closest to the depot.

The number of possible partitions is given by the number of ways to choose $k-1$ items (left of the depot) from among $n-1$ items (the smallest c_j is fixed at k),

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!\,(n-k)!}. \tag{4.3}$$

Our branch and bound scheme builds partial solutions by assigning the c_j in sequence, from the largest to the smallest, to the left or to the right of the depot. And therefore partial solutions are built up by first assigning locations at the ends of the line and then moving in toward the depot. So for example, a partial solution of $(L, L, L, R, R, 0, 0, \ldots, 0)$ indicates that the three largest c_j are assigned left of the depot, the next two to the right of the depot, and the rest of the locations closest to the depot are unassigned or free .

For a particular partial solution, we let f_L be the number of free locations to the left of the depot and f_R be the number of free locations to the right. We let the probability that no pick is required, so far in a partial solution, to the left of the depot, be given by

$$q_L = \prod_{i=0}^{k-1-f_L} q_i, \tag{4.4}$$

and the probability that no pick is required to the right of the depot be given by

$$q_R = \prod_{i=0}^{n-k-f_R} q_{n+1-i}. \tag{4.5}$$

The next theorem establishes that if one side has more free locations remaining than the other, and larger probability of no pick required, then the next largest c_j should be assigned to that side.

Theorem 4.6. *If a partial solution has $f_L \geq f_R$ and $q_L \geq q_R$ then there exists an optimal way to layout the remaining free locations with the next largest c_j assigned to the left side. And conversely, if $f_L \leq f_R$ and $q_L \leq q_R$ then there exists an optimal way to layout the remaining free locations with the next largest c_j assigned to the right side.*

Proof. See Appendix, section A.2. □

Thus Theorem 4.6 shows that c_1 (the largest) should be assigned to the longest side. Corollary 4.5 can be seen to follow from Theorem 4.6, and we will see that the other special case given in Jarvis and McDowell also follows from this theorem. But perhaps more importantly, Theorem 4.6 generalizes the result of Jarvis and McDowell that provides us a computationally useful cut in our branch and bound scheme — allowing us to fathom any partial solution that violates the theorem.

The partitioning aspect of our problem and Theorem 4.4 leads to two natural heuristics. An *Alternating Increasing Layout* (proposed by Jarvis and McDowell) partitions the c_j in increasing (non-decreasing) sequence, alternating between left of the depot and right of the depot. When one side is full the other side is filled with the remaining c_j's. Thus the Alternating Increasing Layout (AIL) begins assigning the locations closest to the depot and then works its way out to the ends of the pick line.

We introduce an *Alternating Decreasing Layout* that partitions the c_j in decreasing (non-increasing) sequence, alternating between left of the depot and right of the depot, to provide another upper bound for our procedure. So the Alternating Decreasing Layout (ADL) begins assigning the locations farthest from the depot and working its way in.

The following result established in Jarvis and McDowell also follows as a consequence of Theorem 4.6 — when k is centrally located then an ADL or AIL is optimal.

Lemma 4.7. *When a single depot is located in the center of the line, then an ADL or AIL layout will minimize* $E[S_k|N > 0]$.

Proof sketch When n is odd and $k = (n+1)/2$, then an ADL and AIL layout are identical. Furthermore, at each iteration of the ADL as the next largest c_j is located, the requirements of Theorem 4.6 are maintained. See Jarvis and McDowell for a different argument. □

So, when k is located centrally or at the start of the line, our branch and bound scheme will return the optimal layout immediately by fathoming any partial solution that violates Theorem 4.6. For other k, the two heuristics, ADL and AIL, are used in our branch and bound scheme for upper bounds. Lower bounds for a partial solution are generated in a simple manner: If c_j is the first (largest) unassigned c_j we then assign $c_j, c_{j+1}, \ldots c_{j+f_L-1}$ values to the remaining free locations to the left of the depot. And we assign $c_j, c_{j+1}, \ldots c_{j+f_R-1}$ values to the free locations to the right of the depot. So the largest unassigned values are used on the remaining free locations on *both* sides of the depot — an infeasible solution that provides a lower bound.

Table 1 provides run times in CPU seconds for instances with $n = 40$ locations. Three different data sets were generated; using a seed $0 < r < 1$, the c_j values are set by $c_1 = 1 - r^n,\ c_2 = 1 - r^{n-1}, \ldots, c_{n-1} = 1 - r^2,\ c_n = 1 - r$. Branching from a partial solution was halted whenever the lower bound was within 1% of the upper bound. The code was written in C and run on a 750 MHz PC. For each problem set, when the depot is near the end of the line (small k) the runs times are favorable because the number of possible partitions is very small — see Equation (4.3). But when the depot is near the center of the line (large k) the

	$n = 40$ Items										
	k										
	1	2	4	6	8	10	12	14	16	18	20
$r = .9$	0.01	0.01	0.03	0.52	5.22	37.71	150.19	395.35	131.30	18.87	< 0.01
$r = .7$	< 0.01	< 0.01	0.09	4.59	93.51	827.52	2973.56	3363.54	680.27	7.09	< 0.01
$r = .5$	< 0.01	< 0.01	0.10	4.08	52.26	235.48	393.81	189.23	24.93	0.46	0.01

Table 1: Run times (in cpu seconds on a 750 MHz PC) to obtain optimal layouts for a single depot at location k with $n = 40$ locations for three different data sets, $r = 0.9, 0.7, 0.5$.

number of possible partitions is the largest; however, the run times are small because Theorem 4.6 provides very effective fathoming when the depot is near the center. However, for intermediate depot locations the algorithm is the least effective as the number of partitions is fairly large and Theorem 4.6 is not as effective. When n is increased beyond 40 the algorithm can still obtain solutions when k is near the end or the center of the line; however some instances with $n = 50$ and intermediate k fail to return an optimal solution after 10 hours of CPU time.

4.2.2 Dual depot item layout

The layout problem for dual depots reduces to a single-depot problem. Since a picker must always traverse between u and v for each order, then the $u - v + 1$ smallest c_j are simply placed between the two depots in any sequence. The remaining problem is then a single-depot layout problem with the remaining c_j to be assigned to the left of u and to the right of v. Thus all our results and algorithms apply to the dual-depot case. We make clear that the solvable special case for a single, centrally located, depot corresponds to the dual-depot case when $u = n - v + 1$, the two depots are equidistant from the center.

4.2.3 No depot item layout

We seek a layout to minimize Equation (3.8),

$$E[T] = E[W|N > 0] + \frac{1}{2}\Big(E[B_L|N_1, N_2 > 0] + E[B_R|N_1, N_2 > 0]\Big).$$

To minimize $E[T]$ directly seems complex. However, what is desired is a compact pick run so that not only $E[W|N > 0]$ is small, but also $\mathrm{E}[B_L|N_1, N_2 > 0]$ and $E[B_R|N_1, N_2 > 0]$. Thus an obvious heuristic is to assume a centrally located single depot and layout the items according to the ADL heuristic. This will minimize $E[W|N > 0]$ about the center of the line, and thus is expected to perform well.

4.3 Single and dual depot full design

To find the optimal way to both layout the items and locate the depot(s) we can simply enumerate the possible depot locations, since we know from Theorems 4.1 and 4.2 that they must be integral. So for a single depot we consider all n possible locations, and then solve for the optimal layout as described in Section 4.2.1. For dual depots, we consider the $n(n-1)/2$ possible dual depot locations, and then for each solve for the optimal layout as covered in Section 4.2.2.

5 Conclusions

Order picking is commonly considered the most critical function in a supply chain; and within the warehouse the order picking function typically accounts for most of the operating costs. Furthermore, new e-commerce operations distinguish themselves from their traditional bricks-and-mortar competition in having to absorb the cost of the order picking function.

Discrete order picking is common due to its simplicity, reliability, and its ability to pick orders quickly upon receipt. However it can also be wasteful, requiring considerable walking per pick. Technology is the first way to help reduce the walking required. The second is through the design — where to locate the depots (if any are used) and how best to layout the product.

We consider three levels of technology: (1) no technology (single-depot configuration); (2) conveyor technology (dual-depot configuration); (3) conveyor and RF technology (no-depot configuration). We find that the effectiveness of each level of technology depends on the characteristics of the order stream — many picks per order or very few. We introduce a new configuration (dual depots with conveyor technology) and find that it performs very well regardless of the type of order stream. Only when the pick density is very small (few picks per order) is the additional investment in RF technology potentially warranted.

So our work provides a manager a model to estimate the impact of different technology investments on his order picking system. And for each level of technology we provide algorithms to optimally design the system — locate depot(s) and layout the product.

Acknowledgments

I would like to thank Gary Eppen, Kevin Gue, and Linus Schrage for helpful discussions and comments on earlier drafts. I would like to thank the management at Urbanfetch.com and Revco Drugs, Inc. for sharing their system and insights. Furthermore, I would like to acknowledge the generous support of the Graduate School of Business at the University of Chicago.

References

[1] J. J. Bartholdi, III, D. D. Eisenstein and R. D. Foley. "Performance of bucket brigades when work is stochastic", *Operations Research* **49**(5) (2001).

[2] R. De Koster, T. Le-Duc, and K. J. Roodbergen. "Design and control of warehouse order picking: a literature review", Working paper (2006). To appear in *European Journal of Operational Research.*

[3] E. P. Chew and L. C. Tang. "Travel time analysis for general item location assignment in a rectangular warehouse", *European Journal of Operational Research* **112** 582-597 (1999).

[4] R. L. Francis, L. F. McGinnis and J. A. White. *Facility Layout and Location: An Analytical Approach,* Prentice-Hall (1992).

[5] E. D. Frazelle. *World-Class Warehousing,* Logistics Resources International, Inc., Atlanta, GA. (1996).

[6] R. W. Hall. "Distance approximations for routing manual pickers in a warehouse", *IIE Transactions* **25**(4) 76-87 (1993).

[7] J. M. Jarvis and E. D. McDowell. "Optimal product layout in an order picking warehouse", *IIE Transactions* **23**(1) 93-102 (1991).

[8] J. O. Matson and J. A. White. "Operational research and material handling", *European Journal of Operational Research* **11** 309-318 (1982).

[9] H. D. Ratliff and A. S. Rosenthal. "Order-picking in a rectangular warehouse: a solvable case of the traveling salesman problem", *Operations Research* **31**(3) 507-521 (1983).

[10] M. B. Rosenwein. "A comparison of heuristics for the problem of batching orders for warehouse selection", *International Journal of Production Research* **34**(3) 657-664 (1996).

[11] J. A. Tompkins and J. D. Smith. *The Warehouse Management Handbook,* Tompkins Press (1998).

[12] J. P. van den Berg. "A literature survey on planning and control of warehousing systems", *IIE Transactions* **31** 751-762 (1999).

A Proof of Theorems 4.6

We will begin with a short lemma and some new notation that will be useful for both of these proofs.

A.1 Some Technical Preliminaries

We will make use of the following simple lemma.

Lemma A.1. *If $x < y$ and $\alpha < \beta$ then $\alpha y + \beta x < \alpha x + \beta y$.*

Proof. $(\alpha x + \beta y) - (\alpha y + \beta x) = (y - x)(\beta - \alpha) > 0$. □

We also will use the following notation when dealing with a sequence $s_1, s_2, \ldots$.

$$\boldsymbol{s}(i,j) = s_i + s_i s_{i+1} + s_i s_{i+1} s_{i+2} + \cdots + (s_i s_{i+1} \ldots s_j) \quad \text{for } i \le j. \tag{A.1}$$

And we let the product of the terms of the sequence from i to j, $i \le j$, be denoted as

$$\dot{\boldsymbol{s}}_{i,j} = s_i s_{i+1} \ldots s_j. \tag{A.2}$$

A.2 Proof of Theorem 4.6

Consider a partial solution with f_L free locations to the left of the depot and f_R free locations to the right. We need only consider the one case in which $f_L \ge f_R$ and $q_L \ge q_R$. We relabel the free locations to the left of the depot, $q_{k-f_L}, q_{k-f_L+1}, \ldots, q_{k-1}$, as $a_1, a_2, \ldots a_{f_L}$ so that a_1 is furthest from the depot and a_{f_L} is closest. We also relabel the free locations to the right of the depot as $b_{f_R}, b_{f_R-1}, \ldots, b_2, b_1$ where b_{f_R} is closest to the depot and b_1 is farthest.

Rewriting Equation (4.2), using Equations (4.4) and (4.5) and the notation of Equations (A.1) and (A.2), we seek to maximize

$$\boldsymbol{q}(1, k-1-f_L) + q_L\, \boldsymbol{a}(1, f_L) + \boldsymbol{q}(n, n-k-f_R) + q_R\, \boldsymbol{b}(1, f_R),$$

and after eliminating the assignments that are fixed we seek to maximize

$$q_L\, \boldsymbol{a}(1, f_L) + q_R\, \boldsymbol{b}(1, f_R). \tag{A.3}$$

Our theorem states that an optimal layout exists when the next largest c_j is assigned to a_1 whenever $f_L \ge f_R$ and $q_L \ge q_R$. Suppose to the contrary that one claims to have completed the layout and maximized Equation (A.3) with a larger c_j assigned to b_1 instead of a_1. We will show that by swapping at least a_1 and b_1 in such a proposed layout that we can obtain a new layout that is no worse.

We consider two cases.

1. $\boldsymbol{a}(1, f_L) \geq \boldsymbol{b}(1, f_R)$ in the proposed layout.

We rewrite Equation (A.3) as

$$q_L\, a_1 + q_L\, a_1\, \boldsymbol{a}(2, f_L) + q_R\, b_1 + q_R\, b_1\, \boldsymbol{b}(2, f_R). \tag{A.4}$$

Since $b_1 > a_1$, we know $\boldsymbol{a}(2, f_L) > \boldsymbol{b}(2, f_R)$, and thus, since $q_L \geq q_R$, we have from Lemma A.1, that Equation (A.4) increases if we swap a_1 and b_1.

2. $\boldsymbol{a}(1, f_L) < \boldsymbol{b}(1, f_R)$ in the proposed layout.

We let j be the largest index so that $\dot{\boldsymbol{a}}_{1,j} < \dot{\boldsymbol{b}}_{1,j}$ (we know such an index exists since $a_1 < b_1$). If $j < f_R$, then we rewrite Equation (A.3) as

$$q_L\, \boldsymbol{a}(1, j) + q_L\, \dot{\boldsymbol{a}}_{1,j}\, \boldsymbol{a}(j+1, f_L) + q_R\, \boldsymbol{b}(1, j)\ + q_R\, \dot{\boldsymbol{b}}_{1,j}\, \boldsymbol{b}(j+1, f_R). \tag{A.5}$$

Since $f_R \leq f_L$, and by our choice of j, we have $\dot{\boldsymbol{b}}_{1,j}\boldsymbol{b}(j+1, f_R) < \dot{\boldsymbol{a}}_{1,j}\boldsymbol{a}(j+1, f_L)$. And thus, $\boldsymbol{b}(j+1, f_R) < \boldsymbol{a}(j+1, f_L)$.

We rewrite our assumption $\boldsymbol{a}(1, f_L) < \boldsymbol{b}(1, f_R)$ as

$$\boldsymbol{a}(1, j) + \dot{\boldsymbol{a}}_{1,j}\boldsymbol{a}(j+1, f_L) < \boldsymbol{b}(1, j) + \dot{\boldsymbol{b}}_{1,j}\boldsymbol{b}(j+1, f_R).$$

But since $\dot{\boldsymbol{b}}_{1,j}\boldsymbol{b}(j+1, f_R) < \dot{\boldsymbol{a}}_{1,j}\boldsymbol{a}(j+1, f_L)$ we have $\boldsymbol{a}(1, j) < \boldsymbol{b}(1, j)$.

So if we swap the j values $b_1, \ldots, b_j$ and $a_1, \ldots, a_j$ in Equation (A.5) we get

$$q_L\, \boldsymbol{b}(1, j) + q_L\, \dot{\boldsymbol{b}}_{1,j}\, \boldsymbol{a}(j+1, f_L) + q_R\, \boldsymbol{a}(1, j)\ + q_R\, \dot{\boldsymbol{a}}_{1,j}\, \boldsymbol{b}(j+1, f_R), \tag{A.6}$$

which, from Lemma A.1 is larger than Equation (A.5).

If however, $j = f_R$ then we can show in a similar, but somewhat simpler, fashion that a swap of all $b_1, \ldots, b_{f_R}$ with $a_1, \ldots, a_{f_R}$ will result in a new layout that is no worse.