Appendix

A. Derivation of Inequality (8)

In the following, we first recall Inequalities (6), (7), and (8), and notation used. Recall that nodes L^0 and L^1 denote the child nodes obtained from tree node L by branching on customer $k \in N$, where L^0 represents the node where customer k does not receive an incentive and L^1 the node where customer k is incentivized. Node L^1 is dominated by L^0 if Inequality (6) holds:

$$\min_{y\in\mathcal{D}(L^1)} \sum_{\omega\in\Omega} P_\omega(y)S_\omega(y) \geq \min_{y\in\mathcal{D}(L^0)} \sum_{\omega\in\Omega} P_\omega(y)S_\omega(y). \tag{6}$$

Let $y^1 \in \arg\min_{y\in\mathcal{D}(L^1)} \sum_{\omega\in\Omega} P_\omega(y)S_\omega(y)$ be a policy that minimizes the left-hand-side of the Inequality (6). To prove the dominance of node L^0, it suffices to show that there exists at least one policy $\tilde{y} \in \mathcal{D}(L^0)$ for which Inequality (7) holds:

$$\sum_{\omega\in\Omega} P_\omega(y^1)S_\omega(y^1) \geq \sum_{\omega\in\Omega} P_\omega(\tilde{y})S_\omega(\tilde{y}). \tag{7}$$

In the following, we prove the existence of such policy $\tilde{y} \in \mathcal{D}(L^0)$ if Inequality (8) holds:

$$d_k \geq F_k(y^1), \text{ where } F_k(y^1) \equiv \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)C_\omega - \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)C_\omega. \tag{8}$$

Proof. To prove the dominance of L^0, it suffices to show that there exists at least one policy $\tilde{y} \in \mathcal{D}(L^0)$ for which Inequality (7) holds. Consider the policy $\tilde{y} \in \mathcal{D}(L^0)$, with $\tilde{y}_i = y_i^1$ for all $i \in N \setminus k$, and $\tilde{y}_k = 0$. Then, Inequality (7) is equivalent to:

$$\sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} P_\omega(y^1)S_\omega(y^1) + \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} P_\omega(y^1)S_\omega(y^1) \geq \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} P_\omega(\tilde{y})S_\omega(\tilde{y}) + \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} P_\omega(\tilde{y})S_\omega(\tilde{y}). \tag{A.1}$$

Using the definitions of $S_\omega(y)$, $P_\omega(y)$, $D_\omega(y)$ and Table 1 from Section 3, we have for policies y^1 and $\tilde{y}$:

$$P_\omega(y^1)S_\omega(y^1) = \begin{cases} (1-\Delta_k)\cdot \prod\limits_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot\left(\sum_{i\in N\setminus k} w_i y_i^1 d_i + C_\omega\right) & \omega\in\Omega, \omega_k = 0 \\ \Delta_k \cdot \prod\limits_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot\left(\sum_{i\in N\setminus k} w_i y_i^1 d_i + d_k + C_\omega\right) & \omega\in\Omega, \omega_k = 1 \end{cases} \tag{A.2}$$

$$P_\omega(\tilde{y})S_\omega(\tilde{y}) = \begin{cases} \prod\limits_{i\in N\setminus k} P_{\omega_i}(\tilde{y}_i)\cdot\left(\sum_{i\in N\setminus k} w_i \tilde{y}_i d_i + C_\omega\right) & \omega\in\Omega, \omega_k = 0 \\ 0 & \omega\in\Omega, \omega_k = 1. \end{cases} \tag{A.3}$$

Substituting Equations (A.3) and (A.2) into Inequality (A.1) yields:

$$\begin{aligned} &(1-\Delta_k)\cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot \sum_{i\in N\setminus k} w_i y_i^1 d_i + (1-\Delta_k)\cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot C_\omega + \\ &\Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot \sum_{i\in N\setminus k} w_i y_i^1 d_i + \Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y_i^1)\cdot (d_k + C_\omega) \geq \\ &\sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(\tilde{y}_i)\cdot \sum_{i\in N\setminus k} w_i \tilde{y}_i d_i + \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(\tilde{y}_i)\cdot C_\omega. \end{aligned} \tag{A.4}$$

By definition of policies $\tilde{y}$ and y^1, it holds that $\tilde{y}_k = 0$, $y^1_k = 1$, and for all $i \in N \setminus k$ we have $\tilde{y}_i = y^1_i$, which implies $P_{\omega_i}(y^1_i) = P_{\omega_i}(\tilde{y}_i)$. Thus, (A.4) can be simplified as:

$$-\Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot \sum_{i\in N\setminus k} w_i y^1_i d_i - \Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \left(\prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot C_\omega \right) + \tag{A.5}$$
$$\Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot \sum_{i\in N\setminus k} w_i y^1_i d_i + \Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \left(\prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot C_\omega \right) + d_k \Delta_k \cdot \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \geq 0.$$

We observe that the expression $\prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot \sum_{i\in N\setminus k} w_i y^1_i d_i$ does not depend on ω_k. Consequently, the first and the third terms in (A.5) are identical and can be eliminated. Moreover, as the probability of all realizations of the final delivery locations for customers $i \in N \setminus k$ add up to 1, it follows that $\sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) = 1$ in the last term of (A.5). As a result, we obtain from (A.5):

$$d_k \geq \sum_{\substack{\omega\in\Omega:\\ \omega_k=0}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot C_\omega - \sum_{\substack{\omega\in\Omega:\\ \omega_k=1}} \prod_{i\in N\setminus k} P_{\omega_i}(y^1_i) \cdot C_\omega,$$

which is equivalent to Inequality (8). Hence, we have proven that if Inequality (8) holds, then there exists at least one policy $\tilde{y} \in \mathcal{D}(L^0)$ (i.e., with $\tilde{y}_i = y^1_i$ for all $i \in N \setminus k$, and $\tilde{y}_k = 0$), which satisfies Inequality (7), implying that node L^0 dominates node L^1. □

B. Cumulative probability of scenarios in set $B^i(k)$ given a policy y

In this appendix, we derive a closed-form expression for the cumulative probability of scenarios in set $B^i(k)$ given a policy y:

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i'\in N\setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}). \tag{B.1}$$

Recall, that we consider a customer k, and a set of all vertices excluding this customer, $V \setminus k$, ordered such that a vertex i precedes a vertex j (denoted $i \prec j$) if $c_{ki} < c_{kj}$ (if $c_{ki} = c_{kj}$, then lexicographic order is followed). For each pickup point $t \in \mathcal{P}$, $N(t)$ denotes the set of customers whose nearest pickup point is t. For a given vertex $i \in V \setminus \{k\}$ set $B^i(k) = \{\omega^{\bar{k}} \in \{0,1\}^{n-1} : i \in V_{\omega^{\bar{k}}}\} \setminus \cup_{j\prec i} B^j(k)$ includes all vectors $\omega^{\bar{k}}$ whose vertex set $V_{\omega^{\bar{k}}}$ includes vertex i but not vertices $\forall j \prec i$.

To compute the probability given in expression (B.1), we consider 4 collectively exhaustive and mutually exclusive cases with respect to the type of vertex $i \in V$ and its position relative to other vertices:

1. The depot precedes vertex i (i.e., $0 \prec i$) or there exists a pickup point $t \in \mathcal{P}$ and a customer $j \in N(t)$, such that $t \prec i$ and $j \prec i$.

   If the depot vertex precedes vertex i, then the set $B^i(k)$ is empty, as no scenarios exist where vertex i is visited but the depot is not. Similarly, if the distance from customer k to a pickup point $t \in \mathcal{P}$ and a customer $j \in N(t)$ is less than that to vertex i, then the set $B^i(k)$ is also empty. This is because customer j needs to be served either at home or at the nearest pickup point; hence, no scenarios exist where vertex i is visited but one of the vertices j or t is not. Consequently, we trivially derive:

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i'\in N\setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}) = 0. \tag{B.2}$$

2. Vertex i is either a depot or a customer, and the conditions of *Case 1* are not satisfied (depot vertex does not precede vertex i, and there is no pickup point $t \in \mathcal{P}$ and a customer $j \in N(t)$ such that $t \prec i$ and $j \prec i$).

   Given that customer choices are independent, the events of visiting each customer are also independent. Similarly, the event of visiting each pickup point is independent of the event of visiting other pickup points. Moreover, by the criteria of *Case 2* there is no pickup point $t \in \mathcal{P}$ and a customer $j \in N(t)$ such that both of them precede vertex i (i.e., $t \prec i$ and $j \prec i$). Therefore the event of visiting any pickup point $t \prec i$ preceding vertex i depends on the choice of customers $j \in N(t)$ none of whom precedes vertex i. As a result, the probability given in Equation (B.1) is defined by *independent* events to not visit the vertices preceding vertex i and the event of visiting vertex i.

   Hence, for the situation when i is a depot vertex, the expression given in Equation (B.1) is equal to the probability to visit the depot (equal to 1, as the depot is always visited), multiplied by the probability that customers preceding vertex i do not require visit (equal to $P_1(y_j)$ for each customer $j \in N$, $j \prec$i), and multiplied by the probability that pickup points preceding vertex i do not require visit (equal to $\prod_{j\in N(t)} P_0(y_j)$ for each pickup point $t \prec i$, as in these scenarios all customers for whom pickup point t is the closest must opt for home delivery):

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i' \in N \setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}) = 1 \cdot \prod_{\substack{j\in N \\ j \prec i}} P_1(y_j) \prod_{\substack{t\in\mathcal{P} \\ t\prec i}} \prod_{j \in N(t)} P_0(y_j). \tag{B.3}$$

   We derive a similar expression for the case when vertex i is a customer, but in this case the probability to visit vertex i is $P_0(y_i)$:

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i' \in N \setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}) = P_0(y_i) \cdot \prod_{\substack{j\in N \\ j \prec i}} P_1(y_j) \prod_{\substack{t\in\mathcal{P} \\ t\prec i}} \prod_{j \in N(t)} P_0(y_j). \tag{B.4}$$

3. Vertex i is a pickup point and there exists a customer $j \in N(i)$ preceding this pickup point ($j \prec i$), and the conditions of the *Case 1* are not satisfied (depot vertex does not precede vertex i, and there is no pickup point $t \in \mathcal{P}$ and a customer $j \in N(t)$ such that $t \prec i$ and $j \prec i$).

   As there exists a customer $j \in N(i)$ preceding pickup point i ($j \prec i$), by definition of set $B^i(k)$, customer j must be served at the pickup point i, which imposes the condition that pickup point i must be visited (i.e., the probability to visit this pickup point is 1 in these scenarios). Thus, the expression given in Equation (B.1) is equal to the probability to visit the pickup point i (equal to 1, as this pickup point must be visited in these scenarios), multiplied by the probability that customers preceding vertex i do not require visit (equal to $P_1(y_j)$), and multiplied by the probability that other pickup points preceding vertex i do not require visit (equal to $\prod_{j\in N(t)} P_0(y_j)$ for each pickup point $t \prec i$):

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i' \in N \setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}) = 1 \cdot \prod_{\substack{j\in N \\ j \prec i}} P_1(y_j) \prod_{\substack{t\in\mathcal{P}\setminus\{i\} \\ t\prec i}} \prod_{j \in N(t)} P_0(y_j). \tag{B.5}$$

4. Vertex i is a pickup point and there is no customer $j \in N(i)$ satisfying $j \prec i$, and the conditions of the *Case 1* are not satisfied (depot vertex does not precede vertex i, and there is no pickup point $t \in \mathcal{P}$

and a customer $j \in N(t)$ such that $t \prec i$ and $j \prec i$).

In this case, the probability that pickup point i is visited depends only on the choice of customers $j \in N(i)$ who are not preceding vertex i, and equals to $1 - \prod_{j \in N(i)} P_0(y_j)$. Similarly to the previous case, we obtain:

$$\sum_{\omega^{\bar{k}} \in B^i(k)} \prod_{i' \in N \setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'}) = (1 - \prod_{j \in N(i)} P_0(y_j)) \cdot \prod_{\substack{j \in N \\ j \prec i}} P_1(y_j) \prod_{\substack{t \in \mathcal{P} \setminus \{i\} \\ t \prec i}} \prod_{j \in N(t)} P_0(y_j). \tag{B.6}$$

Equations (B.2)–(B.6) provide the closed-form expression for the cumulative probability $\sum_{\omega^{\bar{k}} \in B^i(k)} \prod\limits_{i' \in N \setminus k} P_{\omega^{\bar{k}}_{i'}}(y_{i'})$ of scenarios in set $B^i(k)$ given a policy y, for all potential cases.

C. Proof for Dominance Rule 2

In the following, we present the proof for Dominance Rule 2, which is restated below for convenience:

Dominance Rule 2 *Let two customers $j, k \in N$ be located within a distance $c_{jk} \leq \frac{1}{2}\min(d_j, d_k)$. Then, any node that offers an incentive to exactly one of these customers is dominated by a node that offers an incentive to neither of these customers.*

Proof. Consider a node L such that *(i)* the incentive decisions have been made for all customers except for customer k and *(ii)* no incentive is offered to customer j, i.e., $l_j = 0$. By branching on customer k, two child nodes can be obtained from node L. As the child nodes are complete policies, they are leaf nodes, hence, the best policy that can be obtained from each node is itself. Let y^1 is the leaf node where customer k is offered an incentive. Then, to prove that y^1 is dominated by the child node which does not offer an incentive to customer k, it is sufficient to have (by Condition (8) and Equation (11)):

$$d_k \geq F_k(y^1), \text{ where } F_k(y^1) = \sum_{\omega^{\bar{k}} \in \{0,1\}^{n-1}} \prod_{i \in N \setminus k} P_{\omega^{\bar{k}}_i}(y^1_i)(C_{(\omega^{\bar{k}},0)} - C_{(\omega^{\bar{k}},1)}). \tag{C.1}$$

Since $y^1_j = 0$, customer j is certainly visited at home; hence, Inequality (C.1) is equivalent to:

$$d_k \geq \sum_{\substack{\omega^{\bar{k}} \in \{0,1\}^{n-1} \\ \omega^{\bar{k}}_j = 0}} \prod_{i \in N \setminus k} P_{\omega^{\bar{k}}_i}(y^1_i)(C_{(\omega^{\bar{k}},0)} - C_{(\omega^{\bar{k}},1)}). \tag{C.2}$$

For any $\omega^{\bar{k}}$ where customer j is visited at home, it holds that $C_{(\omega^{\bar{k}},0)} - C_{(\omega^{\bar{k}},1)} \leq 2c_{jk}$. Further, as the probability of all realizations of the final delivery locations for customers $i \in N \setminus k$ add up to 1, it follows that $\sum_{\substack{\omega^{\bar{k}} \in \{0,1\}^{n-1}: \\ \omega^{\bar{k}}_j = 0}} \prod\limits_{i \in N \setminus k} P_{\omega^{\bar{k}}_i}(y^1_i) = 1$. As a result, the right hand side of (C.2) has an upper bound of $2c_{jk}$. Therefore, to prove that y^1 is dominated, it is sufficient to have:

$$d_k \geq 2c_{jk}. \tag{C.3}$$

That is, if Condition (C.3) holds, the policy that offers an incentive to customer k but not to customer j is dominated by the policy which offers neither of these customers an incentive. By symmetry, we can prove that the policy that offers an incentive to customer j but not to customer k is dominated by the policy which offers neither of the customers an incentive if:

$$d_j \geq 2c_{jk}. \tag{C.4}$$

Conditions (C.3) and (C.4) together imply that any node that offers an incentive to exactly one of the customers k or j is dominated by a node that offers an incentive to neither of these customers if:

$$\min(d_j, d_k) \geq 2c_{jk}. \tag{C.5}$$

This completes the proof. □

D. Algorithm for calculating the bundle probability

In the following algorithm, we denote the set of all h-size subsets of a given set of pickup points $\tilde{\mathcal{P}}$ with $\{A \subseteq \tilde{\mathcal{P}} : |A| = h\}$. For each pickup point $t \in \mathcal{P}$, $N(t)$ denotes the set of customers whose nearest pickup point is t. A pseudocode of the recursive algorithm for computing bundle probability $\sum_{\omega \in B(\tilde{\mathcal{P}})} \min_{y \in \mathcal{D}(L)} P_\omega(y)$ is given in Algorithm 2.

Algorithm 2: Bundle probability

Input: Subset of pickup points $\tilde{\mathcal{P}}$, node $L = (l_i)_{i \in N}$

Output: bundle probability $\sum_{\omega \in B(\tilde{\mathcal{P}})} \min_{y \in \mathcal{D}(L)} P_\omega(y)$

Function `Probability`($\tilde{\mathcal{P}}, L$)
- $\pi \leftarrow 1 \cdot \prod_{t \in \mathcal{P} \setminus \tilde{\mathcal{P}}} \prod_{i \in N(t):\ l_i \neq 0} (1 - \Delta_i)$
- **for** $h \leftarrow |\tilde{\mathcal{P}}| - 1$ *to* 0 **do**
  - **for** $\mathcal{P}' \in \{A \subseteq \tilde{\mathcal{P}} : |A| = h\}$ **do**
    - $\pi \leftarrow \pi -$`Probability`($\mathcal{P}', L$)
- **return** π

return $\pi =$`Probability`($\tilde{\mathcal{P}}, L$)

As the base case, the algorithm computes the cumulative probability of scenarios, in which only pickup points from a subset of $\tilde{\mathcal{P}}$ are visited. This probability equals the probability that home delivery is selected by all customers $i \in N$ for whom the non-visited pickup points $t \in \mathcal{P} \setminus \tilde{\mathcal{P}}$ are the closest, while the choice of other customers is not restricted:

$$1 \cdot \prod_{t \in \mathcal{P} \setminus \tilde{\mathcal{P}}} \min_{y \in \mathcal{D}(L)} \prod_{i \in N(t)} P_0(y_i) = 1 \cdot \prod_{t \in \mathcal{P} \setminus \tilde{\mathcal{P}}} \prod_{\substack{i \in N(t) \\ l_i = 0,1}} P_0(y_i = l_i) \prod_{\substack{i \in N(t) \\ l_i = \#}} \min_{y_i \in \{0,1\}} P_0(y_i) = 1 \cdot \prod_{t \in \mathcal{P} \setminus \tilde{\mathcal{P}}} \prod_{i \in N(t): l_i \neq 0} (1 - \Delta_i), \tag{D.1}$$

where $\min_{y_i \in \{0,1\}} P_0(y_i)$ is obtained from Table 1 in Section 3.2.

Equation (D.1) includes the probabilities of scenarios, where not all pickup points of set $\tilde{\mathcal{P}}$ are visited. To subtract their probability, in the next step we calculate the bundle probability for all subsets $\mathcal{P}' \subset \tilde{\mathcal{P}}$. Calculating the probability of the bundle $\pi(\mathcal{P}', L)$ is the base case of the algorithm.