Solving Truckload Procurement Auctions Over an Exponential Number of Bundles

Richard Li-Yang Chen, Shervin AhmadBeygi, Amy Cohn, Damian R. Beil, Amitabh Sinha,
Solving Truckload Procurement Auctions Over an Exponential Number of Bundles

Richard Li-Yang Chen, Shervin AhmadBeygi, Amy Cohn
Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109-2117
{richchen@umich.edu, shervin@umich.edu, amycohn@umich.edu}

Damian R. Beil, Amitabh Sinha
Ross School of Business, University of Michigan, Ann Arbor, Michigan 48109-1234
{dbeil@umich.edu, amitabh@umich.edu}

Truckload carriers provide hundreds of billions of dollars worth of services to shippers in the United States alone each year. Internet auctions provide these shippers with a fast and easy way to negotiate potential contracts with a large number of carriers. Combinatorial auctions have the added benefit of allowing multiple lanes to be considered simultaneously in a single auction. This is important because it enables carriers to connect multiple lanes in continuous moves or tours, decreasing the empty mileage that must be driven, and therefore increasing overall efficiency. On the other hand, combinatorial auctions require bidding on an exponential number of bundles to achieve full economies of scope and scale, which is not tractable except for very small auctions. In most real-world auctions, bidding is instead typically limited to a very small subset of the potential bids. We present an implicit bidding approach to combinatorial auctions for truckload procurement that enables the complete set of all possible bids to be considered implicitly, without placing the corresponding burden of an exponential number of bids on the bidders or the auctioneer. We present the models needed to solve this problem. We then provide extensive computational results to demonstrate the tractability of our approach. Finally, we conclude with numerical analysis to assess the quality of the solutions that are generated and to demonstrate the benefits of our approach over existing bidding methods in practice.

Keywords: combinatorial auctions; implicit bidding approach; truckload procurement; network flows

1. Introduction and Motivation

U.S. freight transportation expenditures in 2005 exceeded $700 billion. Of this amount, $300 billion was accounted for by the truckload segment (American Trucking Association 2007). In many corporations, transportation expenditures can be as high as 30% of the overall cost of goods sold (Ballou 1992); furthermore, trucking is often the dominant cost. Therefore, reducing trucking expenditures can greatly reduce a shipper’s cost of goods sold and improve profitability.

Typically, shippers estimate their freight to be shipped in an upcoming year based on the prior year’s shipments (Foster and Strasser 1991). When contracting out truckload services, a shipper puts forth a request for quotes (RFQ) for a network of lanes. Traditionally, carriers (i.e., trucking companies) have submitted quotes for individual lanes in the RFQ. This is akin to a single-item reverse auction, where each lane is awarded independently to a single carrier using a single criterion, usually price (Sheffi 2004).

Today, Internet auctions provide shippers with a fast and easy way to simultaneously negotiate multiple potential contracts with a large number of carriers. The use of the Internet as an auction medium has the benefit of decreasing information gathering, participation, and transaction costs, as well as increasing geographic and temporal conveniences (Lucking-Reiley 2000). Large corporations such as The Home Depot, Walmart Stores, and Staples Inc. rely on applications from software providers to procure billions of dollars worth of services annually via Internet auctions (Elmaghraby and Keskinocak 2002; Ledyard et al. 2002; Caplice and Sheffi 2003). Prominent providers of such software currently include CombineNet and Manhattan Associates.

Over the past decade, many of these auctions have allowed bidders to bid on combinations of lanes instead of bidding only on individual lanes. Such auctions, called combinatorial auctions, have three stages. First, the auctioneer (on behalf of the shipper) announces multiple lanes for bid (henceforth, bid lanes) in the auction. Second, the bidders (here, the carriers) submit bids for sets of bid lanes (bundles), rather than bidding on each bid lane individually. Third, the
auctioneer determines the best set of bundles that collectively cover each bid lane, and awards contracts for these bundles (rather than awarding individual bid lanes) to the corresponding bidders.

An important benefit of combinatorial auctions is that they often make it possible to capture the benefits of substitution effects and complementarities, in which the value of a set is not simply the sum of its parts. Using a combinatorial auction in such cases allows bidders to express their true preferences, with the goal of finding better allocations. This is the case in truckload shipping, because of, for example, the fact that carriers must not only transport the bid lanes that they have been awarded, but must also return drivers home. If carriers can string together multiple loads to form a continuous move (tour), then they can decrease their empty mileage and thereby reduce cost.

We illustrate this in Figure 1 with a simple example with two loads. Here, we see that the bidder’s bid price of transporting a load from A to B is \( l \cdot x \) (the direct movement price per mile \( l \) times the distance \( x \)) to move the load from A to B) plus \( e \cdot x \) (the empty movement price per mile \( e \) times the distance \( x \)) to return the driver home from B to A). The price of transporting a load from B to A is computed similarly. The price of transporting both loads, however, is not \( 2(l \cdot x + e \cdot x) \) but instead only \( 2l \cdot x \), because the two loads can be combined to form a single tour, without any empty movements. If the loads were auctioned individually, in two separate auctions, then the bidders would face the following dilemma when bidding on the first load. If they bid high presuming the full price (including the empty return trip), they might lose the auction for bidding too high a price. But if they bid low presuming only the direct movement price and then did not win the second load, they could lose money. A combinatorial auction ameliorates this by allowing three bids: one for winning only load A, one for winning only load B, and one for winning both.

More broadly, combinatorial auctions allow carriers to bid on bundles of lanes to produce more efficient movements. The efficiency gained in combinatorial bidding, in turn, allows carriers to submit more aggressive bids, thereby reducing transportation costs for the shipper.

There is a substantial stream of literature on the explicit computation of bundle bids for truckload transportation auctions. Wang and Xia (2005), Lee, Kwon, and Ma (2007), and Chang (2009) provide methods for identifying bundles that are likely to be good to bid on, and efficiently computing the bids for those bundles. Filgozzi, Mahmassani, and Jaillet (2007) study a dynamic setting: they show how to compute noncombinatorial bids for entire contracts in an environment where contracts are put up for bid sequentially over time. All of these papers recognize that truckload transportation services is an area where the potential for economies of scale and scope is particularly rich. In fact, there is opportunity to leverage the complementarities in lanes among different shippers even before the auction begins (Ergun, Kuyzu, and Savelsbergh 2007a, b). Likewise, there is also opportunity to leverage synergies in lanes among different carriers to improve efficiencies in the transportation of contracted lanes (Filgozzi, Mahmassani, and Jaillet 2006; Song and Regan 2003; Houghtalen, Ergun, and Sokol 2008; Krajewska and Kopfer 2006).

However, two major hurdles remain that prevent the full realization of the benefits of combinatorial auctions. The first is bid expression and communication: to fully express economies of scale and scope among all items being auctioned, bidders must construct and submit bids for an exponential number of subsets of these items (2\( n \) – 1 for an \( n \)-item auction). This is clearly intractable for all but the smallest instances. The second hurdle is in solving the winner determination problem (WDP), typically formulated as a set partitioning problem (Balas and Padberg 1976), to select the lowest-cost set of bundles such that each item is in exactly one bundle. WDP is an integer program with an exponentially large number of binary variables, and thus also intractable for all but the smallest instances.

De Vries and Vohra (2003) presented the state of knowledge for solving combinatorial auctions and suggest the used of an “oracle” to alleviate the burden of expressing and communicating an exponential number of bids. The auctioneer invokes the appropriate oracle at any stage of an auction to determine the bid for a particular bundle. Alternatively, an auctioneer may specify a bidding language (Rothkopf,
The goal of our research is to show that the underlying structure of a truckload procurement problem can be exploited similarly, enabling us to find solutions to fully enumerated auctions in practical timeframes.

This research extends the existing literature on combinatorial truckload procurement auctions (CTPA). For example, Sheffi (2004) and Caplice and Sheffi (2006) presented the state of knowledge for CTPAs. Ledyard et al. (2002) and Elmaghraby and Keskincocak (2002) described early uses of combinatorial auctions for truckload procurement. Song and Regan (2005) and Raychaudhuri and Veeramani (2006) addressed carrier bidding strategies in multi-round auctions. Guo et al. (2006) extended the carrier assignment models used in WDP to include shipper nonprice objectives and carrier transit point costs.

Nevertheless, the full benefits of combinatorial auctions for truckload procurement have not yet been achieved in practice. One recent study (Plummer 2002) showed that only 28% of carriers submit bids of more than one lane in combinatorial auctions and the majority of these carriers only submit 2–7 multi-lane bundles because of practical constraints on bid preparation time, computational resources, and technical expertise at their disposal.

We propose an implicit bidding approach to truckload procurement auctions that can (implicitly) capture the full, exponential set of bundles. This approach leverages the fact that there is a known and amenable structure underlying the cost of servicing a given set of bid lanes. Specifically, the least-cost tour (or set of tours) needed to cover a set of lanes can be computed by solving a minimum cost flow problem. We therefore propose to embed this underlying cost structure (which we refer to as a bid-generating function (BGF)) directly into WDP. This eliminates the need for the bidder to compute and communicate an exponential number of bids. Furthermore, we will show that the resulting WDP can be reformulated as a multicommodity flow (MCF) problem of polynomial size. Our computational results demonstrate the practical performance of the implicit bidding approach.

The contributions of this research are in:

1. presenting a new implicit bidding approach for CTPA that enables the complete set of all possible bids to be considered implicitly, and thus achieves full economies of scope;

2. developing tractable models to solve a basic truckload procurement auction to optimality, in single round, fully considering (implicitly) the exhaustive set of all possible bids;

3. showing how the power of mathematical programming can enable this basic problem to be extended to include additional important real-world operational considerations; and

4. taking advantage of this new capability to solve fully enumerated truckload procurement auctions as a tool for conducting numerical analysis on the characteristics of CTPA solutions.
The paper is organized as follows. In §2, we formally present combinatorial auctions for truckload procurement. In §3, we introduce the implicit bidding approach for CTPA. In §4, we present computational experiments focusing on the tractability of the implicit bidding approach, solution characteristics under a variety of conditions, and performance comparison to bidding methods in practice. We conclude in §5 with a summary of our contributions and our suggestions for future research.

2. CTPA

In a basic truckload procurement auction, the auctioneer specifies a set of bid lanes, each defined by an origin, a destination, and a volume (typically corresponding to the expected number of loads). Given a bundle of bid lanes, carriers determine the least-cost set of tours to serve these bid lanes, then use this cost in computing their bid price for this particular bundle. For example, in a first-price auction, carriers typically bid true cost plus a percentage-based markup (Song and Regan 2004). Throughout the manuscript, we will assume a first-price auction with a percentage-based markup for the sake of exposition. However, our approach is applicable to other auction mechanisms as well. Finally, the auctioneer solves a WDP to select bundles and allocate the corresponding lanes to the winning carriers.

2.1. Computing Bundle Bids

To understand how carriers compute their bids, we must first understand their cost structure. The carriers’ cost of service can be decomposed by individual movements. The obvious cost is the direct movement cost, associated with actually moving a load from lane origin to destination. This cost is well understood by the carrier and is largely a function of distance (fuel, equipment depreciation, driver’s wage, tolls, etc.).

In addition, there is the repositioning cost associated with moving a truck from the destination of one lane to the origin of the next, so as to form tours. To minimize cost, and thus improve the probability of winning bids, carriers must try to build efficient continuous movements with minimal empty mileage. This can be accomplished not only by combining bid lanes, but also by taking advantage of a carrier’s pre-existing contracted lanes and opportunities on the spot market. For example, Figure 2 shows a sequence of movements for efficiently transporting three bid loads.

These repositioning opportunities are the key to determining the actual bid price for any bundle of lanes to be served, resulting in significant synergies and complementarities, as seen in the example above. Because repositioning opportunities are often not known with certainty at the time of the auction, carriers typically estimate these opportunities for each directed city pair \((i, j)\) in the network. One way to represent this is with an \(n\)-tiered step function, where each tier represents capacity and cost estimates of a different type of repositioning opportunity. For example, one tier might represent the expected number of movements from a preexisting contracted lane (with other shippers), which can be used “for free,” as these movements represent hired rather than empty movements. Another tier might represent the potential for partial connections: preexisting lanes that require the driver to travel empty from \(i\) to some nearby location before picking up the load and delivering it to some location near \(j\), thereby incurring limited empty mileage costs. Estimates of spot market opportunities would be represented by additional tiers as well. Finally, the highest cost tier, with infinite capacity, represents empty movements from \(i\) to \(j\).

Figure 3 provides an example of such a step function. More generally, high-traffic city pairs would have high-capacity, low-cost tiers because of the abundance of backhaul opportunities, while low-traffic city pairs would have lower-capacity, higher-cost tiers representing the decreased likelihood of finding complementary lanes.

Finally, we reiterate that given a set of bid lanes, direct movement cost and repositioning opportunities (costs and capacities), carriers determine the least-cost set of tours to serve these bid lanes, then use this cost in computing the bid price, typically, cost plus a percentage-based markup.
2.2. Traditional WDP
Once the bids have been submitted, the auctioneer then solves the WDP to select bundles and allocate lanes to winning carriers. The traditional winner determination formulation (T-WDP) is as follows:

Sets
- $\mathcal{X}$: set of carriers and $K = |\mathcal{X}|$
- $\mathcal{A}^d$: set of arcs representing bid lanes and $L = |\mathcal{A}^d|$
- $\mathcal{F}_k$: set of bundles submitted by carrier $k$

Parameters
- $D_a$: expected volume of bid lane $a$, $\forall a \in \mathcal{A}^d$
- $\delta_{sk}$: number of lane $a$ movements in bundle $s$ of carrier $k$, $\forall a \in \mathcal{A}^d$, $k \in \mathcal{X}$, $s \in \mathcal{F}_k$
- $b_k^s$: carrier $k$’s bid price (cost plus a percentage-based markup) for bundle $s$, $\forall k \in \mathcal{X}$, $s \in \mathcal{F}_k$

Variable
- $x_s^k$: binary variable that takes value 1 if carrier $k$ is awarded bundle $s$ and 0 otherwise, $\forall k \in \mathcal{X}$, $s \in \mathcal{S}_k$

(T-WDP) min $\sum_{k \in \mathcal{X}} \sum_{s \in \mathcal{F}_k} b^k_s x^k_s$ subject to:

1. $\sum_{k \in \mathcal{X}} \sum_{s \in \mathcal{F}_k} \delta_{sk} x^k_s = D_a \quad \forall a \in \mathcal{A}^d$,
2. $\sum_{s \in \mathcal{F}_k} x^k_s \leq 1 \quad \forall k \in \mathcal{X}$,
3. $x^k_s \in \{0, 1\} \quad \forall k \in \mathcal{X}$, $s \in \mathcal{F}_k$.

The objective (1) is to minimize the total cost the shipper pays for procuring truckload services for all lanes in $\mathcal{A}^d$. Constraint set (2) states that bundles must be chosen such that each lane in $\mathcal{A}^d$ is fully covered by the awarded bundles. In a fully enumerated CTPA, an additional constraint set (3) stating that each carrier can be awarded at most one bundle is imposed (note that each bundle might contain more than one tour). This constraint set is needed to ensure that we do not select a combination of bundles for a given carrier such that, in total, the combination of bundles violates some of the carrier’s operational constraints (for example, using more repositioning capacity than there exists).

We conclude this section by re-iterating the fact that, for practically sized truckload procurement auctions with thousands of lanes, it is, of course, not possible to explicitly enumerate all bundle bids. For each lane $a \in \mathcal{A}^d$, a carrier may bid a volume of zero up to $D_a$, so the total number of distinct combination is equal to

$$\Lambda \equiv \prod_{a \in \mathcal{A}^d} (D_a + 1).$$

Even in the case of a regional carrier who bids only on a subset of a few dozen lanes, this would still entail millions of bundle combinations. Instead, carriers submit bids for only a small subset of the bundles, because of practical constraints on bid preparation time, computational resources, and technical expertise available at their disposal (Plummer 2002). As a result, the solution quality of CTPAs in practice is often compromised.

3. Implicit Truckload Combinatorial Auctions
In the majority of the combinatorial auction literature, bundle prices are assumed to be exogenously endowed. For CTPAs, although most of the literature continues with this exogenous endowment of bundle bids, there is more recent literature on the identification and pricing of bundles. For instance, Wang and Xia (2005) and Lee, Kwon, and Ma (2007) show how carriers can efficiently identify promising bundles to bid on as well as the bid price for those bundles. Figliozzi, Mahmassani, and Jaillet (2007) also explore how synergies impact the pricing of bids for bundles in a dynamic environment, where demand for loads to be served arrives over time.

We follow a similar approach for the purpose of establishing how carriers generate bids for specific bundles. The key idea, as recognized by the papers cited above, is that bundle prices depend critically on the network structure, existing commitments, and potential future commitments of the carrier. We model these by a BGF.

More important, the structure of this BGF can now be exploited using an implicit bidding approach to solve WDP using BGFs directly, in lieu of the actual bids. This enables the exhaustive set of bundles to be considered implicitly without sacrificing tractability.

3.1. The BGF
Given a bundle of bid lanes $s$, carrier $k$’s price to service these lanes is comprised of the individual direct movement and repositioning movement prices. For a given set of lanes, the direct movement prices are fixed and known in advance. The repositioning movement prices, on the other hand, depend on the continuous moves that the carrier constructs to minimize the total price of the bundle. These repositioning moves enable carriers to exploit synergies and complementarities that exist in serving lanes of a bundle. As such, the price of a bundle may be significantly different from the sum of bid prices of individual lanes in that bundle, as discussed in §2.1.

For a given carrier $k$, the problem of determining the least-cost set of continuous moves (and thus the value of the corresponding bids) to serve a set of bid lanes can be computed by solving a network flow problem on a directed graph, $G(\mathcal{N}, \mathcal{A})$. In this graph, node set $\mathcal{N}$ represents origin, destination, and/or transhipment cities. Arc set $\mathcal{A}$ represents direct movement and repositioning movement lanes.
with associated arc prices (cost plus profit markup) and capacities. Repositioning movement arcs may include the carrier’s estimates of opportunities from pre-existing contracted lanes, anticipated opportunities on the spot market, and empty movements. In particular, we construct one arc for each tier of repositional opportunities, one for each tier of a directed city pair (as described in §2.1).

The problem is then to create the least-cost set of tours in this network such that each bid lane in s is covered. The notation and formulation for this BGF (which we denote by $f^k$) are as follows.

### Sets

$\mathcal{N}$ set of nodes corresponding to arc origins or destinations

$\mathcal{A}^k$ set of arcs representing carrier $k$’s estimated repositioning opportunities, one for each tier of a directed city pair

### Parameters

$\theta(a)$ origin of arc $a$, $\forall a \in \mathcal{A}^k \cup \mathcal{A}^d$

$\mathcal{D}(a)$ destination of arc $a$, $\forall a \in \mathcal{A}^k \cup \mathcal{A}^d$

$p_s^k$ carrier $k$’s price for a unit movement on arc $a$, $\forall a \in \mathcal{A}^k \cup \mathcal{A}^d$

$u_s^k$ carrier $k$’s estimated repositioning capacity on arc $a$, $\forall a \in \mathcal{A}^k$

$x^k_s$ a vector of carrier $k$’s bid volume, where element $x^k_s$ represents the bid volume of lane $a \in \mathcal{A}^k$ in bundle $s$.

Observe that the price of a unit movement on arc $a$, denoted $p_s^k$, is strictly a function of the carrier’s cost to complete a movement on this arc plus a profit markup. In turn, the profit markup can account for factors such as competitive strategy, geographic location of depots, transshipment centers, etc. Carriers can also use the parameter $p_s^k$ to indicate an undesirable lane (e.g., a lane outside their region of coverage) by setting a very high value. Last, in the context of pricing bundle $s$, the quantity $x^k_s$ is a parameter of the BGF $f^k$ and not a variable.

### Variable

$y^k_a$ the number of repositioning movements on arc $a$ made by carrier $k$, $\forall a \in \mathcal{A}^k$.

The BGF can now be represented as the following integer program:

\[
\begin{align*}
(BGF) \quad f^k(x^k) & = \min \sum_{a \in \mathcal{A}^k} p_s^k x^k_a + \sum_{a \in \mathcal{A}^d} p_s^d y^k_a \\
\text{subject to} \quad & \sum_{a \in \mathcal{A}^k : \theta(a) = i} x^k_a + \sum_{a \in \mathcal{A}^d : \theta(a) = i} y^k_a = \sum_{a' \in \mathcal{A}^k : \theta(a') = i} x^k_{a'} + \sum_{a' \in \mathcal{A}^d : \theta(a') = i} y^k_{a'} \quad \forall i \in \mathcal{N},
\end{align*}
\]

The objective function (5) states that carrier $k$‘s price for bundle $s$ is the sum of direct movement prices (which depends solely on $s$ and is known in advance) and the repositioning movement prices (which depends on the chosen routing). Because these sums are over distinct sets of movement arcs, the carrier can submit different prices for bid arcs and repositioning arcs over the same origin-destination pairs. Constraint set (6) ensures flow conservation at each node in the network; that is, the number of movements into a node must be equal to the number of movements out of the node. This ensures that the resulting solution is a set of tours covering lanes in $s$. Constraint set (7) states that the repositioning capacity used must be less than or equal to the available capacity.

$f^k(x^k)$ has two important structural characteristics that have significant impact on tractability. First, the integrality restrictions (8) can be replaced by non-negativity constraints $y^k_a \geq 0$, $\forall a \in \mathcal{A}^k$ because the constraint matrix of this problem is totally unimodular. Second, BGF can be reformulated as a circulation problem via a simple variable redefinition. Circulation problems, which are special cases of minimum cost flow problems, are well known to be easy to solve. For examples of polynomial time algorithms for the circulation problem, please refer to Ahuja, Magnanti, and Orlin (1993).

### 3.2. The Implicit Winner Determination Problem (I–WDP)

Using the traditional auction mechanism described in §2, each carrier must solve a BGF (5–8) to obtain a bid price for each bundle of interest. For real-world truckload procurement auctions with thousands of lanes, constructing bid prices for the full exponential set of bundles is not possible. Furthermore, even if carriers could compute and communicate bids for all bundles, the auctioneer could not solve the corresponding exponentially large WDP. We show that these hurdles can be overcome by using an implicit bidding approach, which directly embeds a carrier’s BGF into WDP. The resulting polynomially sized (with respect to bid lanes, number of carriers, and number of nodes) model is solution equivalent to the fully enumerated T–WDP but, in contrast, is tractable for practically sized instances.

The thrust of this implicit bidding approach is the following. Rather than submit an exponential number of bundle price pairs, each carrier $k$ instead submits the parameters of the BGF, $f^k$. These parameters are simply a list of all the arcs with corresponding prices (including any profit markups) and capacities. Note,
of course, that a carrier choosing not to bid on particular lanes (e.g., those outside their geographic region of coverage) would simply not include those arcs in their parameters.

The auctioneer can then imbed $f^k$ directly into WDP, which can be reformulated as

$$\min \sum_{k \in \mathcal{K}} f^k(x^k)$$

subject to:

$$\sum_{k \in \mathcal{K}} x^k_a = D_a \ \forall a \in \mathcal{A}^d,$$  \hspace{1cm} (10)

$$x^k_a \in \mathbb{Z}^+ \ \forall a \in \mathcal{A}^d, \ k \in \mathcal{K}. $$  \hspace{1cm} (11)

Observe that Equations (9)–(11) implicitly capture substitution effects and complementarities, resulting in a fully enumerated truckload procurement auction, where each winner is awarded exactly one bundle (possibly empty). This bundle is described by the single vector of decision vector $x^k$ (with one element per bid lane), taking the place of the set of vectors $x^k$ (with one vector per bundle) previously defined in §3.1. For each $a \in \mathcal{A}^d$, $x^k_a$ is the volume of bid lane $a$ assigned to carrier $k$. Note, however, that in place of $K \cdot A$ binary variables, there are now only $K \cdot L$ integer variables in the model described by (9)–(11). As an example, in an auction with 10 carriers and 100 bid lanes (each with a volume of 10), this translates to a reduction from more than $1.37 \times 10^{10}$ binary variables to only 1,000 integer variables.

Of course, even with this reduction in size, the new formulation may still be quite difficult to solve, depending on the structure of $f^k$. As we have noted in §3.1, however, $f^k$ is simply a circulation problem. Thus, embedding (5)–(8) in place of the function $f^k$ leaves us with the following mixed-integer program, which we denote by I–WDP:

$$(\text{I–WDP}) \min \sum_{k \in \mathcal{K}} \left[ \sum_{a \in \mathcal{A}^d} p^k_a x^k_a + \sum_{a \in \mathcal{A}^d} p^k_a y^k_a \right]$$

subject to:

$$\sum_{k \in \mathcal{K}} x^k_a = D_a \ \forall a \in \mathcal{A}^d,$$  \hspace{1cm} (13)

$$x^k_a + \sum_{a \in \mathcal{A}^d: \ell \in \mathcal{A}(a)} y^k_a
\quad \quad = \sum_{a \in \mathcal{A}^d: \ell \in \mathcal{A}(a)} x^k_{a'} + \sum_{a \in \mathcal{A}^d: \ell \in \mathcal{A}(a)} y^k_{a'} \quad \forall i \in \mathcal{N}, \ k \in \mathcal{K}, $$  \hspace{1cm} (14)

$$y^k_a \leq u^k_a \ \forall k \in \mathcal{K}, \ a \in \mathcal{A}^d,$$  \hspace{1cm} (15)

$$x^k_a \in \mathbb{Z}^+ \ \forall k \in \mathcal{K}, \ a \in \mathcal{A}^d,$$  \hspace{1cm} (16)

$$y^k_a \in \mathbb{Z}^+ \ \forall k \in \mathcal{K}, \ a \in \mathcal{A}^d\quad \forall i \in \mathcal{N}, \ k \in \mathcal{K}. $$  \hspace{1cm} (17)

Note that $x$ is no longer a fixed parameter in I–WDP, but now a vector of decision variables. I–WDP has two sets of variables, $x$ and $y$, representing bid lane assignments and the usage of carrier’s repositioning capacities, respectively. The objective function (12) minimizes the total price attributed to direct movements and repositioning movements. Lane cover constraint set (13) stipulates that all bid lanes must be covered by selected carriers. Constraint set (14) ensures flow conservation of nodes for each carrier; that is, the number of movements into a node must be equal to the number of movements out of the node, thereby ensuring that the resulting allocation defines a set of continuous moves (tours). Constraint set (15) states that the repositioning capacities used to complete the tours must be less than or equal to the capacities available.

So, rather than solve an exponentially sized (with respect to the number of bid lanes, number of carriers, and number of nodes) T–WDP, we can instead solve a polynomially sized I–WDP. I–WDP is solution equivalent to a fully enumerated T–WDP, as formally stated by Proposition 1 (see Appendix A). Finally, as proven by Proposition 2 (see Appendix A), we observe that this formulation is a special case of the MCF problem. Although theoretically difficult (Even, Itai, and Shamir 1976), multicmodity flow problem is known to be easy to solve in practice for many real-world instances (Ahuja, Magnanti, and Orlin 1993). This is the case for the truckload procurement auctions, as we will demonstrate through computational results in §4.1.

An implicit assumption in our model is that bidders (carriers) can easily compute the $p^k_a$ values. While sometimes computing $p^k_a$ may be as straightforward as adding variable costs (fuel, wages, etc.), amortized fixed costs (equipment depreciation, for example) and a profit margin, carriers may wish to incorporate other considerations such as competition on the lane and expectations of future traffic. We refer the reader to Caplice and Sheffi (2006) for a discussion on how carriers price individual lanes. There is also a stream of literature (e.g., Figliozzi, Mahmassani, and Jaillet 2007; Song and Regan 2005) on pricing bundles of lanes, in static or dynamic environments; these approaches also provide insight on pricing single lanes. Our work only requires bids prices to be computed for single lanes rather than multilane bundles, and then uses the implicit bidding approach to obtain a WDP that is substantially easier to solve.

3.3. Operational Considerations

In addition to the basic constraints shown in (12)–(17), shippers and carriers may have other operational considerations to take into account. Although all models are, of course, simplifications of the real world, we expand our formulation to capture some of the most natural operational considerations. We provide several examples as follows.
First, we begin by defining auxiliary variables \( q^k_a \) and \( q^k \) for use in defining constraints corresponding to these operational considerations. Let \( q^k_a \) be a binary variable that takes value 1 if carrier \( k \) is awarded at least one load in lane \( a \) and 0 otherwise, and \( q^k \) be a binary variable that takes value 1 if carrier \( k \) is awarded at least one load in any lane and 0 otherwise. The following relationships are then helpful in adding operational constraints:

\[
q^k_a \leq x^k_a \leq D_a q^k \quad \forall a \in \mathcal{A}^x, k \in \mathcal{K}, \tag{18}
\]

\[
x^k_a \leq D_a q^k \quad \forall a \in \mathcal{A}^x, k \in \mathcal{K}, \tag{19}
\]

\[
q^k \leq \sum_{a \in \mathcal{A}^x} x^k_a \quad \forall k \in \mathcal{K}. \tag{20}
\]

**Load volume:** The shipper and carriers may want to restrict the load volume (across all bid lanes) that a carrier can be awarded. A minimum load volume ensures that carriers are awarded at least a threshold volume. A maximum load volume ensures that carriers’ transportation capacities are observed and there is a manageable number of shipper-carrier relationships.

These operational considerations can be modeled as follows, where \( q^k_a \) is the minimum number of loads carrier \( k \) must win (or nothing), \( \bar{\alpha}^k \) is the maximum number of loads carrier \( k \) can win.

\[
q^k_a q^k \leq \sum_{a \in \mathcal{A}^x} x^k_a \leq \bar{\alpha}^k \quad \forall k \in \mathcal{K}. \tag{21}
\]

Constraints (21) say that the total volume of loads awarded to carrier \( k \) must be between \( q^k_a \) and \( \bar{\alpha}^k \) (inclusive) or zero.

**Number of assigned carriers:** The shipper might prefer to award bid lanes to no fewer than \( \beta \) carriers and to no more than \( \bar{\beta} \) carriers, thus ensuring a manageable number of vendor relationships and adequate spreading of risk. We can model such restrictions as follows:

\[
\beta \leq \sum_{k \in \mathcal{K}} q^k \leq \bar{\beta}. \tag{22}
\]

Constraints (22) say that the total number of assigned carriers must be between \( \beta \) and \( \bar{\beta} \) (inclusive).

**Number of assigned carriers per lane:** The shipper may prefer to award a lane to no fewer than \( \gamma_a \) carriers and no more than \( \bar{\gamma}_a \) carriers to ensure operational efficiency.

\[
\gamma_a \leq \sum_{k \in \mathcal{K}} q^k_a \leq \bar{\gamma}_a \quad \forall a \in \mathcal{A}^x. \tag{23}
\]

Constraints (23) say that for each lane, \( a \in \mathcal{A}^x \), the total number of carriers assigned must be between \( \gamma_a \) and \( \bar{\gamma}_a \) (inclusive).

**Favoring of incumbents and performance measures:** There is a cost to the shipper to start a new relationship with a carrier. In practice, incumbents are often favored by 3%–5%—especially for service-critical or time-sensitive lanes (Caplice and Sheffi 2003). Similarly, the level of service (such as percentage on time, claims performance, acceptance rate, etc.) provided by a carrier can also be taken into consideration. These operational considerations can be accounted for by simply adjusting a new carrier’s price coefficients by a constant or multiplicative factor.

### 3.4. Privacy Issues of Implicit Bidding Approach

We have tacitly presumed that carriers would be willing to submit separate price bids for each possible bundle, and that it is not a privacy concern that prevents them from doing so, but rather a practical one—it simply is not tractable to compute and submit such a large number of bids. However, submitting price bids for every possible bundle—explicitly through enumerative bidding or implicitly through a BGF—transmits a substantial amount of information to the auctioneer. Sharing such a large amount of information may naturally raise privacy concerns for the carriers. Broadly speaking, there will always be tension between the perceived risk of providing information versus the opportunities to be gained by leveraging synergies (which can benefit both the carriers and the auctioneer).

Thus, it is worth noting that our implicit bidding approach will be subject to such tensions. For example, transmitting tiered pricing for different repositioning arc volumes in the BGF implicitly reveals capacity information about the carrier. On the other hand, there are substantial gains to be made by an efficient allocation of lanes to carriers, resulting in a win-win situation for both auctioneer and carriers. For the auctioneer, the gains are clear in that the overall price of procuring the transportation services will be lower. For the carriers, the efficient allocation resulting from our approach means that overall, the empty movement by carriers is much lower than in an inefficient allocation. Therefore, carriers can better serve lanes that they win (i.e., serve with lower empty/wasteful repositioning movement), and thereby use their excess capacity to earn more revenues from other markets. While it is, of course, true that some individual carriers may earn lower profit from this approach when compared to some other, less expressive approach, our approach drives overall inefficiencies resulting from unprofitable repositioning moves out of the system. From a longer-term perspective, greater truckload efficiency makes the truckload market more competitive versus alternative shipment modes like rail and air, which has implications for the long-term viability of carriers.

We believe, given the benefits of achieving full economies of scope in a fully enumerated CTPA,
that carriers and shipper have significant incentives to overcome these concerns. Consider an analogous example, vendor managed inventory (VMI) systems (Chopra and Meindl 2007), in which a retailer provides its suppliers with direct visibility to inventory levels. Like our proposed approach, VMI raises concerns about information privacy and security. However, VMI systems are widely used today by large corporations such as Walmart and The Home Depot because of the substantial benefits they provide (Lee, Kwon, and Ma 2007). The practice of collaborative planning, forecasting, and replenishment in the manufacturing and distribution sectors requires even more extensive sharing of information (Aviv 2004). The reason different agents (manufacturers, distributors, retailers, etc.) participate in these systems despite concerns about sharing critical information about capabilities and forecasts is that these systems allow all parties to benefit from the realized efficiencies.

In practice, third-party service providers such as Manhattan Associates and Ariba can provide services such as bidder prequalification and transaction confidentiality to improve information security and privacy and to limit the risk of information leakage. Additionally, emerging research on cryptographically secured auctions (Kudo 1998; Franklin and Reiter 1996) provides an additional way to protect information. We believe our proposed method provides significant incentives for its use and as such may galvanize deployment of existing, or development of new, infrastructures.

4. Computational Experiments

We conducted a set of computational experiments to assess the overall effectiveness of the implicit bidding approach for CTPAs. Specifically, we have focused on:

(1) the tractability of I–WDP and the impact of instance size (number of bid lanes and load volume) on solution time;

(2) the impact of operational constraints (load volume, assigned carriers, assigned carriers per lane) on solution time;

(3) the impact of instance characteristics (repositioning capacity and network structure) on solution characteristics (solution time, number of assigned carriers, and empty movement ratio); and

(4) a comparison to bidding methods in practice in terms of solution times and allocation costs.

Computational experiments were conducted on a Sun x4600-M2 with 8 AMD Opteron 8,218 processors and 64 GB of RAM. The test machine was running Red Hat Enterprise Linux 4. Models and algorithms were coded using C++ and ILOG Concert Technology and solved using ILOG CPLEX 10.0. Parameter files for all computational instances in §4 are available online (Cohn 2009).

4.1. Tractability of I–WDP

We evaluate the performance of I–WDP on randomly generated instances representing various sized auctions. Random instances are controlled by the following parameters: number of nodes (cities), number of carriers, number of bid lanes, repositioning capacity per carrier, and carriers’ price structures (represented by pairs of direct movement and empty movement price per mile).

We generated five sets of experiments representing auctions of size 1,000, 2,000, 3,000, 4,000, and 5,000 bid lanes on a network with 100 nodes representing the 100 most populous cities in the United States. There are 50 carriers (bidders) bidding in each auction. For each set of auctions, we randomly generated 10 instances and report cumulative statistics. The volume of each bid lane is selected uniformly between 50 and 200 loads. A carrier’s repositioning capacity is represented by a set of capacitated, preexisting contracted lanes (that can be used for “free”) and a set of uncapacitated empty movement lanes. The number of preexisting contracted lanes per carrier is selected uniformly between 5% and 15% of the number of bid lanes, with each preexisting contracted lane volume selected uniformly between 10 and 100 loads.

Carriers’ movement prices are generated by multiplying travel distance and a per mile movement price to ensure triangle inequality is satisfied. A carrier’s price to serve an additional load in a bid lane is equal to the distance from bid lane origin to bid lane destination times the carrier’s direct movement price per mile, generated using a Normal distribution, N(1.10, 0.05\(^2\)). Similarly, a carrier’s price to move empty between any city pair is equal to the distance between the city pair times the carrier’s empty movement price per mile generated using a Normal distribution, N(0.80, 0.05\(^2\)).

Results

Solution characteristics for the five auction sizes are shown in Table 1. The median times reported are substantially lower than the averages, which indicates that average solution times are skewed by a few long running instances. Note that, generally, average solution times are inversely proportional to the size (number of bid lanes) of the auction. All else being equal, increasing the number of bid lanes in the auction actually improves solution time. Intuitively, given a fixed-sized network with uniformly distributed lanes, increasing the number of lanes in the auction improves the probability of finding complementary lanes. Therefore, for large auctions, the price of the auction is primarily dominated by direct movement prices and the majority of bid lanes are allocated to a smaller number of low-cost carriers.

The solution times suggest that we can, in fact, solve to optimality fully enumerated CTPAs of up to
5,000 bid lanes (more than 600,000 bid loads) with relative ease. We contrast this again with the traditional approach, which would require each carrier to compute and submit an exponential number of bundle bids and the auctioneer to solve a T–WDP with a corresponding number of binary variables—clearly, an intractable task.

4.2. Impact of Operational Considerations
We next consider the impact on performance of imposing constraints on load volume, number of assigned carriers, and number of assigned carriers per lane. We again consider auctions with 1,000, 2,000, 3,000, 4,000, and 5,000 bid lanes, using the basic instances generated in §4.1 as the baseline. In addition, we conducted the following sets of experiments: basic problem with additional constraints on load volume (constraints (19–21)), basic problem with additional constraints on the number of assigned carriers (constraints (19–20, 22)), and basic problem with constraints on the number of assigned carriers per lane (constraints (18, 23)). We constrained the total load volume awarded to any carrier to be between 50% and 40% of the total load volume. The number of assigned carriers was constrained to be at least 5 and at most 20. Last, we constrained the number of assigned carriers per lane to be at most 10.

Results
As expected, solution time increased with additional constraints. These operational considerations depend on imposing constraints (18) and (19) to define the auxiliary variables $q^a$ and $q^e$. Constraint set (18) and (19) are generally very weak because in most cases, the entire lane volume is not assigned to a single carrier. These types of “big M” constraints typically lead to a weak linear programming (LP) relaxation, which is well known to be computationally undesirable (Nemhauser and Wolsey 1999). There is potential here for future research to develop “stronger” alternative formulations and cutting plane algorithms for improving LP relaxations of constrained I–WDP.

Nonetheless, constrained I–WDPs remain tractable for auctions with up to 5,000 bid lanes (more than 600,000 bid loads), in some cases with improving tractability as the number of bid lanes in the auction grows. The results represented in Table 2 were obtained using default CPLEX solver settings and no preprocessing routines. This also leads us to believe that further improvements in solution times of I–WDP with operational constraints are attainable.

4.3. Impact of Instance Characteristics
In the preceding sections, we demonstrated the viability of the implicit bidding approach by presenting computational results for CTPAs with up to 5,000 bid lanes. Furthermore, we showed that these models can be extended to account for some practical considerations and still maintain tractability.

Now that we have a tractable way to solve, in a single round, fully enumerated CTPAs to optimality (which was not possible in the past), we can also conduct numerical analysis to better understand the performance and characteristics of practical CTPAs. To our knowledge, in the literature, there has not appeared such a study of fully enumerated CTPA outcomes.

In particular, we consider the following two questions:
(1) How does the number of lanes affect solution characteristics?
(2) How do differences in network structure affect solution characteristics?

4.3.1. Effects of Repositioning Lane Quantity.
We first consider how varying carriers’ repositioning capacities impacts solution time, the number of assigned carriers, and empty movement ratio,

<table>
<thead>
<tr>
<th>No. of bid lanes</th>
<th>Avg. no. of bid loads</th>
<th>Avg. repos. capacity (loads)</th>
<th>Solution time (sec.)</th>
<th>Avg.</th>
<th>Median</th>
<th>Std. dev.</th>
<th>Avg. no. of carriers assigned</th>
<th>Avg. no. of B&amp;B nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000</td>
<td>124,425</td>
<td>276,780</td>
<td>181</td>
<td>91</td>
<td>183</td>
<td>6</td>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td>2,000</td>
<td>249,665</td>
<td>551,792</td>
<td>140</td>
<td>99</td>
<td>137</td>
<td>2</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>3,000</td>
<td>376,280</td>
<td>820,569</td>
<td>123</td>
<td>99</td>
<td>93</td>
<td>5</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>4,000</td>
<td>500,493</td>
<td>1,078,834</td>
<td>147</td>
<td>89</td>
<td>123</td>
<td>10</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>5,000</td>
<td>623,654</td>
<td>1,370,068</td>
<td>127</td>
<td>81</td>
<td>122</td>
<td>2</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Average Solution Times for I–WDP with Constraints on Total Load Volume, Number of Assigned Carriers, and Number of Assigned Carriers Per Lane
defined as the ratio of empty movement distance to total (direct plus empty) movement distance. In this numerical experiment, we again consider auctions of 1,000, 2,000, 3,000, 4,000, and 5,000 bid lanes. With the exception of the number of repositioning lanes, parameter settings are identical to those described in §4.1. For each of these auction sizes, we hold the number of bid lanes constant and vary the number of repositioning lanes (of each carrier) as a percentage of the number of bid lanes, ranging from zero to 100%.

Results
In Figures 4–6, we present results for auctions with repositioning capacities for each carrier varying from zero to 50% of the number of bid lanes. Computational results showed that these trends continue to hold beyond these ranges up to 100% of the number of bid lanes.

Our computational results show that CTPAs have special properties at two extremes: when carriers have no repositioning capacities or very large repositioning capacities. At these extremes, the I–WDP is extremely tractable as evident by the small solution times in Figure 4. Furthermore, at these extremes, the number of assigned carriers is relatively small (Figure 5). Intuitively, when carriers have no repositioning capacity, the cost of the auction is dominated by the carriers’ direct and empty movement costs, therefore, a small number of low-cost carriers typically win the majority of the bid lanes. As carriers’ repositioning capacities increase, the likelihood of finding cost-effective connections also increases, leading to decreases in empty movements (Figure 6). As this happens, the majority of continuous movements are formed by combining bid lanes with carriers’ preexisting contracted lanes. In this case, the cost of the auctions is dominated by direct movement cost and a small number of carriers with the lowest direct movement costs again typically win out.

4.3.2. Effects of Network Structure. Next, we evaluate how solution characteristics change as the structure of the network varies. Specifically, what is the impact on solution time, number of assigned carriers, and empty movement ratio? We again consider a network with 100 nodes, which we now divide into six regions. Each bidder is defined to be a national or regional carrier. National carriers have repositioning capacities that are uniformly dispersed throughout the entire network, while regional carriers have repositioning capacities that are concentrated in one specific region. Bid lanes are generated in a uniform network, where bid lanes have randomly selected origin and destination cities, or in a hub-and-spoke network, where bid lanes originate from one of three
Detroit Hub-and-spoke network Uniform network

Figure 7 A Hub-and-Spoke Network (with a Single Hub in Detroit) and a Uniform Network

Notes. Bid lanes in a hub-and-spoke network originate from the hub. Bid lanes in a uniform network are uniformly dispersed throughout the network.

hubs (selected a priori) and terminate at a random node in an adjacent region. Examples of a hub-and-spoke network and a uniform network are provided in Figure 7.

The computational results presented below are based on auctions with a total load volume of 100,000 and 50 carriers. Carriers are assigned 100 repositioning (preexisting contracted) lanes, each with volume 100. We compare computational results for the following four network structures:

1. Network One consists of national carriers bidding on 1,000 bid lanes (with average lane volume of 100) uniformly generated on the network. For each carrier, we uniformly generate 100 repositioning lanes on the network.

2. Network Two consists of regional carriers bidding on 1,000 bid lanes uniformly generated on the network. For each carrier, we randomly generate 100 repositioning lanes within the carrier’s preassigned region.

3. Network Three consists of national carriers bidding on 50 bid lanes (with average lane volume of 2,000) generated on a hub-and-spoke network. For each carrier, we uniformly generate 100 repositioning lanes on the network.

4. Network Four consists of regional carriers bidding on 50 bid lanes (with average lane volume of 2,000) generated in a hub-and-spoke network. For each carrier, we randomly generate repositioning lanes within the carrier’s preassigned region.

Results

Figure 8(a) shows that in a very unstructured network (Network One), with national carriers and uniform bid lanes, the average solution time is 202 seconds. In contrast, with a very structured network (Network Four) consisting of only regional carriers and hub-and-spoke bid lanes, the average solution time is only 13 seconds. In Network One, there is significantly more fractionality; carrier characteristics, in terms of costs and repositioning capacity, are very homogenous and bid lanes are uniformly generated throughout the network. In Network Four, there is less fractionality as carrier characteristics are more heterogeneous; each regional carrier has repositioning capacity that is concentrated in a specific region of the network. Observe that the computational results presented earlier in §§4.1 and 4.2 are based on the least tractable setup, with national carriers and uniform bid lane generation.
lanes. As such, we can expect computational performance of our approach to be even better in real-world networks with some structure.

Figure 8(b) shows that CTPAs on networks with hub-and-spoke bid lanes (Networks Three and Four) result in higher empty movements. This is as expected, because carriers must return empty to hubs (bid lane origin) more often to pick up a bid load. With respect to carrier types, national carriers can better exploit complementarities between their repositioning capacities and bid lanes, and hence service bid lanes more efficiently (0.175 empty movement fraction for national carriers compared to 0.304 empty movement fraction for regional carriers).

With respect to the number of carriers assigned (Figure 8(c)) and number of carriers assigned per lane (Figure 8(d)), less structure implies fewer carriers assigned and more structure implies more carriers assigned. On less structured networks, a few of the lower-cost carriers typically dominate, while on a more structured network, the unique set of repositioning lanes that each carrier brings to the auction plays a key role in forming efficient movements, and so more carriers are likely to be allocated lanes.

4.4. Comparison to Bidding Methods in Practice
The research outlined in this paper is premised on two key ideas. First, the solution quality in a CTPA can improve significantly as the number of combinations bid on grows. Second, the runtime in a CTPA can worsen significantly as the number of combinations bid on grows. In this section, we focused on reconciling these two conflicting issues. We now focus on a comparison between bidding methods in practice and our implicit bidding approach.

4.4.1. Enumerative Bidding Approaches for Low Cardinality Bundles. We use this section to show the impact on solution quality and runtime of combinatorial bidding. A detailed study (Plummer 2002) of carriers’ bidding behavior showed that in practice, only 28% of carriers participating in a combinatorial auction submit multilane bids. Furthermore, among those carriers submitting multilane bids, most submit only two to seven multilane bids of low cardinality (with a median and a mode of two lanes per bundle). In the following experiments, we show that the transition from single-lane bidding to bidding on bundles with even just two or three lanes can dramatically increase solution quality, but also lead to prohibitive increases in runtime. We contrast this with our approach, which can exhaustively (implicitly) consider all bundles of any size, while maintaining tractability for the same problem instances.

For all of the experiments in this section, we assume a small auction in which five carriers bid for lanes (each with volume one) randomly generated in a network of 100 cities. Each carrier has 10 repositioning lanes (each with volume one) that can be used as part of a continuous move for free. Each carrier’s price structure is randomly generated as described in §4.1.

We begin by considering an instance in which 100 bid lanes are being auctioned off by the shipper. We compute the outcome of this auction for five cases:

Case 1. This case permits only single-bid lanes with empty returns. In this case, all bid lanes will be allocated to the lowest price (with respect to direct and empty movement prices) carrier(s).

Case 2. This case permits only single-bid lanes (see Figure 9). However, we allow these bids to reflect the opportunity each carrier has for efficiencies associated with using repositioning lanes in their network for backhaul. Specifically, for each bid lane \( a \in \mathcal{A} \), each carrier computes their lowest price for transporting that lane, taking into consideration the option to use repositioning lanes, and submits the corresponding bid for lane \( a \). In addition, to account for the fact that a carrier cannot use a repositioning lane (with volume one) in more than one bundle, we impose a constraint such that at most one bundle, among those that share a repositioning lane, can be chosen. Finally, we must also include one bid for each lane associated with returning empty, so as to ensure feasibility.

Case 3. This case permits carriers to combine two bid lanes together whenever they create an efficient continuous move (see Figure 10). Specifically, in addition to the bids from Case 2, carriers also bid on each pair of bids lanes, finding the cheapest continuous move that covers both of these lanes (again, using
Table 3 Comparison of Explicit Bidding Approaches to the Implicit Bidding Approach

<table>
<thead>
<tr>
<th>Case</th>
<th>Number of bundles</th>
<th>Min. cost flow (sec.)</th>
<th>T-WDP (sec.)</th>
<th>Total (sec.)</th>
<th>Total cost</th>
<th>Cost differential (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singles — empty returns</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>216,785</td>
<td>51.23</td>
</tr>
<tr>
<td>Singles</td>
<td>600</td>
<td>21</td>
<td>2</td>
<td>23</td>
<td>192,331</td>
<td>34.17</td>
</tr>
<tr>
<td>Singles + doubles</td>
<td>25,350</td>
<td>976</td>
<td>12</td>
<td>988</td>
<td>149,469</td>
<td>4.27</td>
</tr>
<tr>
<td>Singles + doubles + triples</td>
<td>833,850</td>
<td>31,923</td>
<td>681</td>
<td>32,604</td>
<td>145,212</td>
<td>1.30</td>
</tr>
<tr>
<td>I-WDP</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>4</td>
<td>143,349</td>
<td>—</td>
</tr>
</tbody>
</table>

repositioning lanes as well whenever beneficial), and submitting this bid price for the pair of lanes. In addition, to account for the fact that a carrier cannot use a repositioning lane in more than one bundle, we impose a constraint such that at most one bundle, among those that share a repositioning lane, can be chosen.

Case 4. This case extends Case 2 by also allowing all sets of three bid lanes to be combined in continuous moves (see Figure 10). In addition, to account for the fact that a carrier cannot use a repositioning lane in more than one bundle, we impose a constraint such that at most one bundle, among those that share a repositioning lane, can be chosen.

Case 5. This case uses the implicit bidding approach to consider all possible bids of any number of bid lanes.

Table 3 shows the results of these auctions. We present the average results for 10 randomly generated instances. Column one shows the case, column two shows the number of bundles that was considered, and column three shows the total time spent pricing bundles by solving individual minimum cost flow problems. Column four shows the time spent solving T–WDP to determine the allocation of bundles to carriers. Column five shows the total time to solve the model. Column six shows the total cost. The final column gives the ratio of overall cost to the optimal cost to be found if all bundles, regardless of size, are considered. Of course, this is, in fact, the solution obtained by solving the I–WDP.

Observe that moving from single- to two-bid lanes improves the solution quality dramatically (29.90%). The reduction from two- to three-bid lanes (2.97%) is less dramatic but still significant. Although it would appear that moving from three-bid lanes to all lanes does not improve solution quality by a large margin, a 1.30% improvement is still quite meaningful in the trucking industry, where profit margins range from 2% to 4% (Costello 2003). Perhaps more important, these cost improvements associated with moving from single- to two- to three-bid lane bidding come at a substantial cost in terms of runtime, from 23 to 988 to 32,604 seconds (in contrast with the 4 second runtime for I–WDP using the implicit bidding approach). Furthermore, this example is small in size. As the number of bid lanes increases, runtime increases prohibitively, as seen in the next experiment.

Figure 11 projects runtimes for generating all two- and three-bid lane bundles as a function of the number of bid lanes. These were computed by extrapolating from the results in Table 3. On average, each minimum cost flow problem takes 0.038 seconds. We multiply 0.038 seconds by the total number of possible two- and three-bid lane bundles to project runtimes. These projections are actually underestimates, for three reasons. First, as the networks increase in size, the individual minimum cost flow problems take
longer to solve (i.e., the runtime per bid increases). Second, if we consider lanes with volume greater than one load, then the number of possible two- and three-bid lane bundles is even larger. Third, as the number of bids grows very large, computational performance will be impacted by computer memory limitations. Finally, we observe that in the examples shown here, the runtime challenge is a function of the bid generation, with the actual WDP solving quite quickly. This will not always remain true—as the number of bids grows, the impact on IP performance will begin to show.

4.4.2. Selective Bidding Approaches for Low Cardinality Bundles. The results in §4.4.1 suggest that even bidding on all bundles of size three will not be tractable for most truckload auctions, thereby missing substantial opportunities for cost efficiency. A logical counterargument is to suggest not bidding on all bundles of multiple lanes but only bidding on “good” bundles. However, this idea presents two major challenges. First, what defines a “good” bundle? Second, how can we find these “good” bundles?

To address these questions, we conducted the following experiments. We identified three metrics that could be used to evaluate bundles, which are:

1. **Absolute empty mileage**—quality of a bundle is measured by the total empty miles traveled. Lower absolute empty mileage is more desirable.

2. **Empty mileage ratio**—quality of a bundle is measured by the ratio of empty miles traveled to total miles traveled. Lower ratio of empty mileage to total mileage is more desirable.

3. **Random**—bundles are randomly selected from the set of all single, two-, and three-bid lane bundles.

For each of these metrics, we ran a separate auction. In each case, we included all single-lane bids (both with and without the use of repositioning lanes), as in Case 2 of §4.4.1. In addition, we enumerated all two- and three-bid lane bundles and, for each carrier, selected the best bids relative to the metric and included these in the auction. We then solved the auction and reported the final cost. Figure 12 shows the outcomes for each of these three metrics, as the number of bids per carrier varies. These are reported as percentages relative to the optimal value, computed by our implicit bidding approach.

Observe, first, that the optimal solution relative to bidding the exhaustive set of two- and three-bid lane bundles provides a natural lower bound, which, in turn, has a nonzero gap relative to the optimal—in other words, the optimal solution contains some bundles with four or more bid lanes. Note also that although metrics 1 and 2 appear to dominate metric 3 for all but small numbers of bids, it is not always obvious which metric would lead to the better solution.

Furthermore, we see substantial improvement as the number of included bids increases. It is interesting to observe that this does not reflect an increase in the number of multilane bids being included in the final solution. Given 100 bid lanes, for example, at most 33 bundles of size three could ever be included in the final solution. Rather, what we are observing is the fact that the most desirable bundles are not those that best satisfy the given metrics. The reason for this is that we are not concerned with bundles in isolation, but rather with how they fit together with other bundles to complete the auction. As an extreme case, suppose that two carriers shared a common bid lane in their “best” bundles. Because that lane can only be awarded once, only one of those bids could be chosen. Had one of the carriers bid their “second best” bundle, not containing that lane, both bids might be chosen.

Finally, we note that although the runtime for the T–WDP of these instances is certainly shorter than when including all double- and triple-load bid bundles, because of the decrease in size of the IP, we do not know of any efficient way to find these bundles—in our case, we resorted to enumeration, which is no faster than the runtime of Case 3.

**Observations**

We conclude this section by summarizing our observations. First, including multilane bundles, even of small size, can greatly improve solution quality, but at a tremendous impact on runtime. Even small auctions become intractable when including just three-bid lane bundles, never mind bundles of larger cardinality. Second, we cannot overcome this by including only those bundles that are “good bundles”—both because of the computational challenges associated with identifying these bundles and also because the best bundles, in isolation, might not form the best combination of bids. Finally, we observe that our computational results here are only from auctions of a limited size. Although the implicit bidding approach enables us to solve much larger auctions to optimality, we cannot compare the
results to the traditional approach, as the traditional approach cannot be solved to optimality (even across all two- and three-bid lane bundles), except for very small instances. However, as the number of lanes in the auction grows, there may be substantial new opportunities for combining larger sets of lanes (i.e., bundles of size four or more) to incur improved savings, suggesting even more improvement in solution quality as the number of lanes grows.

5. Conclusions and Future Research

In this paper, we introduced an implicit bidding approach to solve CTPAs in a single round, while implicitly considering the exhaustive set of all possible bundles. This approach directly addresses the two main challenges of combinatorial auctions: bidding on an exponentially large set of bundles and solving the corresponding exponentially large WDP. Using the implicit bidding approach, instead of submitting an exponential number of bundles, each carrier simply submits a BGF, which is embedded directly into the WDP. We showed that in truckload transportation, a carrier’s BGF is a circulation problem and the resulting I–WDP is a MCF problem, which is generally known to be tractable in practice. Tractability was demonstrated through extensive computational experiments for auctions with up to 5,000 bid lanes and more than 600,000 loads. Furthermore, I–WDP can be extended to include additional operational considerations while preserving tractability. In short, we presented a new approach and models for solving CTPAs to optimality that are computationally efficient, consider the exhaustive set bundles, and achieve full economies of scope, which is not possible with current approaches.

We also took advantage of this new capability to solve fully enumerated CTPAs to optimality as a tool for conducting numerical analysis on the quality and characteristics of solutions. We showed that, using this approach, shippers can conduct numerical experiments to assess how CTPA characteristics (e.g., solution time, number of carriers assigned, empty movement ratio, etc.) are likely to change with important problem parameters (e.g., number of carriers, number of lanes, carriers’ repositioning capacities, etc.). Additionally, the shipper can use our approach to perform what-if analysis to assess the cost impact of imposing various operational constraints before finalizing contracting decisions. Last, we compared the implicit bidding approach to bidding methods in practice and showed its benefits both in terms of solution quality and runtime.

In terms of future work, we envision two types of research. First, extensions are possible for our work in CTPAs. For instance, additional operational considerations could be addressed, such as regional coverage requirements, backup carrier bids, and maximum tour length constraints. Maximum tour length constraints are applicable because drivers and equipment must be returned home within a limited time window. This constraint set can only be addressed with explicit knowledge of bid lane allocations and the tours constructed to cover these allocated lanes. We are currently addressing this problem using column generation to solve a tour-based model, where each variable represents a viable tour or set of tours. Furthermore, now that we can solve CTPAs in a reasonable amount of time, we can use this tool to assess the quality of various auction mechanisms for truckload procurement. Specifically, how would different auction mechanisms (first price, second price, etc.) perform under various procurement settings? Even with just a single item, revenue—or cost—equivalence between standard auction formats fails if bidders are asymmetric.

Additionally, uncertainties in the cost parameters exist because of spot market variability, carriers’ uncertainties about their existing and future networks, and timing effects; detailed modeling of such uncertainties and development of appropriate solution approaches are interesting, but challenging, directions for future research.

Second, future work could extend the use of the implicit bidding approach to other application domains. Of particular interest is the identification of domains for which the bid generating approach appears amenable. A sample of potential domains include wireless spectrum auctions, energy auctions, and procurement auctions with capacity-constrained suppliers.

Acknowledgments

The authors thank the associate editor and the three anonymous referees for providing helpful and extensive comments that greatly improved this paper, and the Tauber Institute for Global Operations at the University of Michigan for initiating this project via their summer team internship program. This research was supported by the National Science Foundation under Grant CMMI-SEE 0620153 and the STIET Fellowship Program at the University of Michigan.

Appendix. Propositions

PROPOSITION 1. Consider an auction with a set of carriers, bidding for a set of bid lanes \( \mathcal{A} \). For each carrier \( k \in \mathcal{K} \), if the price of a specific bundle \( s \) is given by the solution to \( f^k \) (defined by (5)–(8)), then I–WDP is solution equivalent to a fully enumerated T–WDP.

Proof. We will show that an optimal solution to T–WDP is a feasible solution to I–WDP, with equivalent cost, and vice versa. Let \( \mathcal{J} \) be the set of bundles that correspond to an optimal solution for T–WDP and let \( z_{T-WDP}(\mathcal{J}) \) be the
total cost. Let \((x^*, y^*)\) be the set of vectors that correspond to an optimal solution for I–WDP and let \(z_{L\text{-WDP}}(x^*, y^*)\) be the total cost. □

**Claim 1.** \(z_{L\text{-WDP}}(x^*, y^*) \geq z_{I\text{-WDP}}(x^*, y^*)\)

Proof. For each \(s_i \in \mathcal{S}\), define \(x^i\) to be a vector of size \(L\), where \(x^i_k = w_{a_i}\) (the bid volume of lane \(a_i\) in \(s_i\)) \(a_i \in \mathcal{A}\). Since \(p^i_k\) (price of bundle \(s_i \in \mathcal{S}\)) is obtained by solving \(f_k(x^i)\), there exist vectors \(y^i\) corresponding to the minimum price set of repositioning moves used in \(s_i\). Observe that \((x^i, y^i)\) satisfies constraints (14)–(17) of I–WDP. If we let \((x, y)\) be defined as the concatenation of \((x^i, y^i)\) \(\forall s_i \in \mathcal{S}\), then \((x, y)\) also satisfies (13) and is a feasible solution to I–WDP. Because the cost coefficients of \(f_k\) \(\forall k \in \mathcal{K}\) and I–WDP are identical, \(z_{I\text{-WDP}}(x^i, y^i) = z_{L\text{-WDP}}(x^i, y^i)\). Finally, the optimal solution of I–WDP can only be better, thus we must have \(z_{L\text{-WDP}}(x^*, y^*) = z_{L\text{-WDP}}(x, y) \geq z_{I\text{-WDP}}(x^*, y^*)\). □

**Claim 2.** \(z_{L\text{-WDP}}(x^*, y^*) \leq z_{I\text{-WDP}}(x^*, y^*)\)

Proof. \((x^*, y^*)\) can be decomposed into \((x^i, y^i)\) for each carrier \(k \in \mathcal{K}\). If \(p^i_k(x^*, y^*)\) is the total price to serve the bundle defined by \(x^i\) using repositioning movements corresponding to \(y^i\), then \(z_{L\text{-WDP}}(x^*, y^*) = \sum_{k \in \mathcal{K}} p^i_k(x^i, y^i)\). Observe that repositioning movements \(y^i\) satisfies (6)–(8), and thus \((x^i, y^i)\) is a feasible solution of \(f_k\). Because the optimal solution of \(f_k\) can only be better, we must have \(f_k(x^i) \leq p^i_k(x^i, y^i)\) \(\forall k \in \mathcal{K}\). This implies \(z_{L\text{-WDP}}(x^*, y^*) = \sum_{k \in \mathcal{K}} f_k(x^i, y^i) \leq \sum_{k \in \mathcal{K}} p^i_k(x^i, y^i) = z_{L\text{-WDP}}(x^*, y^*)\). Claims 1 and 2 together imply that \(z_{L\text{-WDP}}(x^*, y^*) = z_{I\text{-WDP}}(x^*, y^*)\). □

**Proposition 2.** I–WDP can be reformulated as a MCF problem.

Proof. The proof is by construction. For each bid lane \(a \in \mathcal{A}\) with corresponding origin \(i\) and destination \(j\), define \(D_{ij} \equiv D_a\) and let \(\mathcal{A}_{ij}\) represent this set of movements. For each carrier \(k \in \mathcal{K}\) and bid lane \(a \in \mathcal{A}\) with corresponding origin \(i\) and destination \(j\), define an arc \((i, j)\) with per unit cost \(c_{ij}^a \equiv p^a_i\); lower bound of \(l_{ij}^a = 0\) and upper bound of \(u_{ij}^a = D_{ij}\). For each carrier \(k \in \mathcal{K}\) and repositioning lane \(a \in \mathcal{A}\) with corresponding origin \(i\) and destination \(j\), define an arc \((i, j)\) with per unit cost \(c_{ij}^a \equiv p^a_k\); lower bound of \(l_{ij}^a = 0\) and upper bound of \(u_{ij}^a = D_{ij}\). Let \(I_k\) represent this set of arcs. Finally, let \(\mathcal{A}_{ij} \equiv \mathcal{A}_{ij} \cup I_k\). Letting \(x^i_k\) represent the amount of flow of commodity (i.e., carrier) \(k \in \mathcal{K}\) on arc \((i, j) \in \mathcal{A}\), I–WDP can be written as

\[
\sum_{k \in \mathcal{K}} \sum_{(i, j) \in \mathcal{A}_{ij}} c_{ij}^a x^i_k + y^i_k \tag{24}
\]

subject to:

\[
D_{ij} \leq \sum_{k \in \mathcal{K}} x^i_k \leq D_{ij} \quad \forall (i, j) \in \mathcal{A}_{ij}, \tag{25}
\]

\[
\sum_{k \in \mathcal{K}} \sum_{(i, j) \in \mathcal{A}_{ij}} x^i_k - \sum_{k \in \mathcal{K}} \sum_{(j, i) \in \mathcal{A}_{ij}} x^i_k = 0 \quad \forall k \in \mathcal{K}, \quad i \in \mathcal{N}, \tag{26}
\]

\[
l_{ij}^a \leq x^i_k \leq u_{ij}^a \quad \forall k \in \mathcal{K}, \quad (i, j) \in \mathcal{A}. \tag{27}
\]

(24)–(27) is a MCF problem. □

**References**


Chen et al.: Solving Truckload Procurement Auctions Over an Exponential Number of Bundles
Transportation Science 43(4), pp. 493–510, © 2009 INFORMS


