Bid behaviour and interactions with the pipeline[*]

Jacob Cosman

June 20, 2019

Introduction   This document describes an approach to modelling users' bidding and acceptance decisions to loads. It creates a mapping between a set of measurable parameters and distributions on one hand and predictions regarding users' response to posted loads and the price pipeline on the other hand. This model is useful for the following purposes:

- Real-time predictions of time to match and probability of successful match.

- Development of an algorithm to dynamically update the pipeline with greater automation.

- An improved understanding of the marginal benefit of an additional app user on load match rate and margins.

- An improved long-run understanding of the impact of changing outside opportunities on carrier behaviour.

- More fully informed real-time load acceptance decisions.

This document outlines a comprehensive and tractable framework for modelling bid and load acceptance behaviour and describes the process for inferring structural parameter values from the data. However, it does not actually provide estimates for parameter values.

Why this matters   Understanding users' bid behaviour and its relationship to the price pipeline offers several potential benefits to Convoy:

1. Higher and more consistent revenue to Convoy by ensuring loads are matched with carriers at the lowest possible cost point.

2. Marginal cost savings from fewer broker interactions per load by decreasing the probability of high-effort attempts to secure a carrier.

3. A more positive user experience by ensuring loads more consistently match through a predictable in-app process.

Improving the response to app users' bidding and load acceptance behaviour (optimally in a way that effectively addresses unmatched loads with low remaining time) could have substantial impacts on Convoy's margins.

Overview of theoretical model   This theoretical model tries to capture the environment and decision set of a carrier using the Convoy app. Accordingly, it excludes information that Convoy knows (e.g. the planned pipeline escalation, the number of other bidders, and the length of time the offer has been online) and includes information that Convoy does not know (e.g. the cost to a carrier of fulfilling a specific load and the time-varying availability of outside opportunities). The model represents the unknown information in a very abstract way although in the future Convoy might have a better sense of carrier- and time-specific costs and outside opportunities.

The model follows Arcidiacono et al. (2016) and other empirical industrial organization literature by situating all decisions in the model environment in continuous rather than discrete time. Rather than all information and decisions happening in discrete chunks of time t, $t+1$, and so on, time is a continuous variable and new information or opportunities to make decisions arrive according to Poisson processes. This requires some additional notation but generates several advantages:

- Thorough use of available high-frequency information, especially app telemetry data.

- A realistic approximation to actual users' experience intermittently checking the app.

- Lower computational intensity (because the probability of two exactly simultaneous events in continuous time is zero).

As with any economic model this approach includes some abstractions from the real world. Some user decisions (e.g. strategically bidding on multiple loads) lie outside the model and some optimal economic behaviour (e.g. fully rational inference of competitive intensity) is not computationally feasible (and arguably not realistic). However, overall the model should provide a reasonable framework for modelling bid behaviour and load acceptance.

A stylized model of a shipment attempt   In this model the unit of observation is a single shipment attempt. A focal app user is making bid and acceptance decisions taking into account Convoy's behaviour (which the user regards as stochastic but not rational) and other app users (which the user regards as rational). That is, the focal user's actions represent the best response to the actions of other users and those other users' actions represent the best response to the actions of the focal user.

The state of the shipment attempt is captured by two variables. The first is the current outstanding offer posted by Convoy in the app (denoted by b). This value is generated by the pipeline and increases over time. The second is the current level of effort exerted by the broker (denoted by e). The effort level ranges from fully automated to additional broker attention to increase the bid to posting on load boards to directly contacting carriers. It corresponds closely to the present level of urgency associated with the shipment attempt. These variables are assumed to be mutual common knowledge. The bid is very clearly visible and presumably users will notice if e.g. the load is posted on load boards (although this latter assumption may be way less justified).

As mentioned previously, the model regards all events as occurring in continuous time. Changes to b and e as well as actions by the model participants arrive according to Poisson processes. The following table describes the events that can occur during a shipment attempt:

Event	Rate	Description of change
Offer update	$\eta(e)$	Offer increases by amount drawn from G
Broker effort increase	$\theta(e)$	Broker effort e increments
Order cancellation	$\delta(e)$	Load removed from Convoy app
Bid opportunity	$\lambda(e)$	User has opportunity to submit bid
Outside opportunity	ϕ	Outside opportunity drawn from F

Convoy's behaviour   In this model Convoy behaves stochastically. Its behaviour is a series of random processes which do not respond strategically to the app users' behaviour. This is arguably a reasonable approximation for the following reasons:

- Convoy's decision process is not visible and its decisions are effectively random from the perspective of the app user.

- Convoy does not currently respond to users' app behaviour in a strategically optimal way but instead uses a series of heuristics.

- It allows for the estimation of the distribution of Convoy's actions at high precision using the large available data set.

Convoy increases its match effort level from e to $e+1$ at rate $\theta(r)$, raises its offer from b to Δb at rate $\eta(e)$ (where Δb itself is drawn from distribution $G(\cdot, e)$), and cancels the order entirely at rate $\delta(e)$. All of these processes depend on the current broker effort level e but not directly on the behaviour of the users. The users regard all three events as occurring at stochastic intervals that depend only on the current visible level of broker effort.

The process of assignment to a load is also stochastic from the user's point of view. It can vary with the user's quality level q. In the Convoy app loads are matched by one of two mechanisms:

- A user clicks to accept the posted offer b. In this case let $\pi_o(e, q)$ be the probability the user is successfully assigned the load. (Hopefully $\pi_o(e) \approx 1$.)

- A user submits a bid B and Convoy accepts that bid. In this case let $\pi(B; b, e, q)$ be the probability that the user is successfully assigned the load. (This is probably close to 0 for $B << b$ and close to 1 for $B \lessgtr b$.)

The user's problem   Each app user comes into the model with a cost of carrying the load c and a quality level q drawn from cumulative distribution function $H(\cdot, \cdot)$. The distribution of user costs H is mutual common knowledge, the realization c is private to the user, and the realization q is known to the user and to Convoy but not to other users. For convenience, let the H_c denote the marginal

distribution of c and H_q denote the marginal distribution of q. The user also has prior beliefs about the number of other users competing for the same load and beliefs about the strategy of other users (both discussed in further detail below). At any time the user also observes the current offer from Convoy b as well as the current level of effort expended by Convoy e. Conditional on all of this the focal app user wants to find a bidding strategy $B(b, e; c, q)$ that maximizes the expected profit associated with this load.

Other users and equilibrium The focal user formulates their strategy in response to their beliefs about the strategies of other users interested in the same load. The focal user assumes that every other user considering the same load is operating according to the bidding strategy $\tilde{B}(b, e)$ with respect to Convoy's current offer b and current broker level e. When modelling this process we will impose the constraint that $\tilde{B}(b, e) = \int_0^\infty \int_0^\infty B(b, e; c, q)\, dH(c, q)$ — that is, that everyone's strategy is the best response to everyone else's strategy. This is almost a weak perfect Bayesian equilibrium. However, it departs somewhat from the formal definition because users do not have fully rationally informed expectations at every moment about the number of competitors.

Number of competitors The focal user in this model does not know how many other users are contemplating bids on the same load. (Convoy, on the other hand, does know exactly how many users are interested — and can use this information to fine-tune the offer pipeline.) Instead, the user will use the available information to form beliefs about the number of competitors n it is facing. In a fully rational and fully informed model the focal user would update beliefs about both the number of other users and the distribution of their quality levels in real time taking into account the expected arrival rate of competitors. However, because this is nether computationally feasible nor a realistic reflection of how users interact with the app, users will have a heuristic that infers the number of competitors based on the probability that no one else has accepted the load yet.

Let $\bar{\rho}(n)$ be the user's prior beliefs that n other users are considering the same load. (For example, these could be the long-run average number of users considering a similar load.) Then, the user will form expectations regarding the number of competitors for this load $\bar{\rho}(n)$ as follows:

$$\rho(n) = \frac{\left[\int_0^\infty \left(1 - \pi\left(\tilde{B}(b, e); b, e; \tilde{q}\right)\right) dH_q(\tilde{q})\right]^n \bar{\rho}(n)}{\sum_{m=0}^\infty \left[\int_0^\infty \left(1 - \pi\left(\tilde{B}(b, e); b, e; \tilde{q}\right)\right) dH_q(\tilde{q})\right]^m \bar{\rho}(m)}$$

Intuitively, if the offer b is low and the effort level e is low and everyone else's prospective bid $\tilde{B}(b, e)$ is very far from acceptance at Convoy then the relatively uninformative initial beliefs $\bar{\rho}$ dominate. However, as b and e rise, a load that is still unfulfilled increases the focal user's weight on the possibility that few or no other users are interested in this load. While this approach does not strictly use all available information as in a fully rational agent model it does capture the correct dynamics.

Value function formulation   Combining all of this generates a function that expresses the value of an offer from the point of view of a user who currently faces an opportunity to bid. This is a relatively complex expression so I will first write out the entire value function and then discuss its components one at a time. The value function V_o for a user currently facing the opportunity to bid is as follows:

$$V_o\left(b, e; c, q\right) = \max \left\{ \underbrace{\pi_o\left(e, q\right)\left(b - c\right) + \left(1 - \pi_o\left(e, q\right)\right) V_w\left(b, e; c, q\right)}_{\text{accept current offer}}, \right.$$

$$\underbrace{\max_{B} \left\{\pi\left(B; b, e, q\right)\left(B - c\right) + \left(1 - \pi\left(B; b, e, q\right)\right) V_w\left(b, e; c, q\right)\right\}}_{\text{submit alternate bid } B},$$

$$\left. V_w\left(b, e; c, q\right) \right\}$$

The user at this point faces three options: they can accept Convoy's current offer, they can submit an alternate bid B, or they can wait until the next opportunity. The first two options are described in further detail below. The value function V_w for a user in between opportunities to bid is as follows:

$$V_w\left(b, e; c, q\right) =$$

$$\underbrace{\left[\eta\left(e\right) + \theta\left(e\right) + \delta\left(e\right) + \phi + \left(1 + \sum_{n=0}^{\infty} \rho_n \int_0^{\infty} \left(1 - \pi\left(\tilde{B}\left(b, e\right); b, e; \tilde{q}\right)\right) \mathrm{d}H_q\left(\tilde{q}\right)\right)\lambda\right]^{-1}}_{\text{discount factor}} \times$$

$$\left[\underbrace{\eta\left(e\right) \int_0^{\infty} V_w\left(b + \Delta b, e; c, q\right) \mathrm{d}G\left(\Delta b, e\right) + \theta\left(e\right) V_w\left(b, e + 1; c, q\right) + \phi\mathbb{E}\left[\max\left\{v, V_w\left(b, e; c, q\right)\right\}\right]}_{\text{non-terminal state changes}} + \right.$$

$$\left.\underbrace{\lambda V_o\left(b, e; c, q\right)}_{\text{next bid opportunity}}\right]$$

    Accepting the current offer   If the user accepts Convoy's current offer then they win this load with probability $\pi_o\left(e, q\right)$ and return to waiting for the next opportunity otherwise.

    Submitting an alternate bid   If the user does not accept Convoy's current offer but instead submits some alternate bid $B > b$ then they win this load with probability $\pi\left(B; b, e, q\right)$ and return to waiting for the next opportunity otherwise.

    Discount factor   Strictly speaking this term isn't a discount factor but it does reflect how quickly the state evolves away from the current state. It includes the rates for new Convoy offers, higher broker effort, removal of the load from the app, a new move opportunity for the focal user, and a successful bid by a user other than the focal user. The higher these terms are, the lower the

value of waiting in any given state. For example, if the rate of removal events $\delta(e)$ is very high, then the value of waiting $V_w(b, e; c, q)$ will be very low, and the user will be less willing to wait if an appealing outside option arises.

Non-terminal changes to state This term encompasses the value of the three state changes that do not end in the offer being removed from the app: an increase in the Convoy offered price (which shifts the state from $V_w(b, e; c, q)$ to $V_w(b + \Delta b, e; c, q)$), an increase in broker effort (which shifts the state from $V_w(b, e; c, q)$ to $V_w(b, e + 1; c, q)$), or the arrival of an outside option. In the first two cases the user continues to wait for the next opportunity. In the latter case the user accepts the outside offer (and stops watching this load) if $v > V_w(b, e; c, q)$ or continues to watch this load if $v \leq V_w(b, e; c, q)$.

Next move opportunity If the next Poisson process to fire is the focal user's move opportunity then the focal user is presented with the move opportunity problem with value function $V_o(b, e; c, q)$ outlined above.

Mapping this to the data This theoretical model describes app users' behaviour as a function of underlying primitive parameters and distributions. It remains to actually measure these. The underlying primitives based on Convoy data are relatively straightforward to capture whereas the primitives based on user behaviour require structural econometric modelling techniques. Strictly speaking, we could roll absolutely everything into one structural econometric model and estimate the parameters that way. However, that approach would be much more computationally intensive and a reducing the dimensionality of the structural estimation will generate more precise estimates.

All of the following parameters can be estimated nonparametrically from internal Convoy data: the offer update rate $\eta(e)$ and offer update size distribution G, the escalation of broker effort $\theta(e)$, the rate at which Convoy cancels offers $\delta(e)$, the bid opportunity arrival rate $\lambda(e)$, and the bid acceptance functions $\pi(B; b, e, q)$ and $\pi_o(e, q)$. Given the depth of available data it should be feasible to estimate all of these as time-adjusted lane-specific averages. These rates depend on the current level of broker effort e as well as the Convoy's level of user quality q so both of these would need to be discretized. For tractability it is probably optimal to restrict these to a few values. The following table describes possible levels for e and q which would be observable in the data:

	e			q	
Level	**Interpretation**		**Level**	**Interpretation**	
0	Automatic pipeline		**0**	High falloff user	
1	Manual pipeline adjustment		**1**	Ordinary user	
2	Posting on load boards		**2**	User with QAU	
3	Calls to carriers		**3**	Elite user	

For the user's prior beliefs regarding the number of competitors $\bar{\rho}(n)$ we can reasonably use a time-adjusted average of the number of users looking at similar offers on the same lane (minus

one because we are looking for competitors to the focal user). Realistically, for many lane/time combinations, this will be very close to zero. In situations where the number of users is higher it may be necessary to develop more refined beliefs.

The remaining attributes (the arrival rate ϕ and distribution F for outside opportunities as well as the joint distribution of user quality and user-specific costs H) are not visible in Convoy data. (However, the marginal distribution H_q is visible in the data.) To estimate these parameters we will need to take the structural model literally and find underlying parameter values which most closely rationalize the observed behaviour as a weak perfect Bayesian equilibrium (with the limited-information beliefs described above). Given the persistent unobserved heterogeneity in the user cost c it is probably optimal to use a nested pseudo likelihood statistical model following Aguirregabiria and Mira (2007). They describe an estimation scheme that recovers structural parameter values while directly enforcing the constraint that all users' behaviour is a best response to other users on the platform. These parameters could potentially depend on outside information — for example, the distribution of c could depend on fuel prices and holiday dates whereas the distribution of v could depend on measures of aggregate regional freight volume as well as the number of offers available on Convoy.

Next steps   The next step for this process is refining the model as much as possible to accurately reflect reality and building parameterizations for the relevant distribution functions. After that, it should be possible to actually estimate this model and validate its behaviour compared to actual observed bid data. If the model provides a reasonable approximation to reality then this will provide an opportunity to predict new pipeline and offer management algorithms that could proceed to A/B testing.

References

Aguirregabiria, Victor and Pedro Mira, "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, January 2007, *75* (1), 1–53.

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson, "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition," *The Review of Economic Studies*, March 2016, *83* (3), 889–931.