Pre-Analysis Plan: Measuring the demand for public transport in Lagos, Nigeria

Daniel Björkegren* Alice Duhaut† Geetika Nagpal‡ Nick Tsivanidis§

August 10, 2023

Abstract

This project will measure demand for public transport in Lagos. It has two components. The first study consists of a field experiment to measure participants value of wait time. The experiment hinges around an SMS-based app (playable on all cellphones) in which participants arrive at a bus stop and are offered a payment amount to wait for a number of minutes before boarding their bus. Enumerators stationed at the bus stops have a tablet which displays a secret code that changes each minute; participants text these codes to our SMS service when they arrive to register their arrival time and receive their offer. If they accept the offer, they send the code displayed after the specified amount of time has elapsed to verify they waited. We randomize the offers of payments and waits; participants' decisions to accept or reject the offers reveal their value of wait time. The second study consists of a field experiment that measures the price elasticity of demand for transport services. In collaboration with the transit regulator in Lagos, we have developed a way in which price subsidies can be administered via the swipe cards used to the pay for the formal bus system in Lagos.

1 Background

Cities in low income countries play a crucial role in structural transformation, and already account for a majority of non-agricultural GDP. But their contribution to growth is being limited by traffic congestion. Congestion is a spatial friction, which distorts interactions between firms and workers. Moreover, transportation is one of the most prominent factors driving air pollution and greenhouse gas emissions in large cities. As more people in low- and middle-income countries move into urban areas, access to efficient public transportation will play a key role both in limiting carbon emissions and reducing the amount of pollution people living in cities are exposed to. However, in many cities in sub-Saharan Africa, more people commute using personal vehicles or informal transportation services than using formal public transit.

*Brown University. Email: dan@bjorkegren.com
†World Bank. Email: aduhaut@worldbank.org
‡Brown University. Email: geetika_nagpal@brown.edu.
§University of California at Berkeley. Email: ntsivanidis@berkeley.edu.

Cities are considering a menu of investments to ease these frictions. What are the impacts from and returns to government investments in new transit systems in Africa? Could similar impacts be obtained by cheaper options, such as regulation of the informal sector? A first part of this project will address the first question by conducting a quasi-experimental impact evaluation of the rollout of 820 new buses across 28 lines in Lagos state combined with three RCTs aimed to causally identify the demand and supply sides of the market. The second part of the project will address the second question by combining these RCTs with a new census of the informal transit network in Lagos and a structural model that characterizes optimal policy for its regulation. Our results will provide evidence-based insights into the relative returns to alternatives for reducing spatial frictions in low income countries.

The remainder of this analysis plan details the two RCTs to measure price and time elasticities of demand, to be extended later for the supply side RCT before it is run.

2 Price RCT

Researchers are partnering with the Lagos Metropolitan Area Transit Authority on a randomized evaluation to study the impacts of introducing a formal public bus system (BRI) on commuters.

Researchers will conduct a household recruitment of approximately 2800 commuters from areas near transit to participate in the commuter intervention, 2000 of which will participate in the "Price Experiment" and 800 of which will participate in the "Wait Time Experiment".

In the Price Experiment, commuters will be randomly assigned to one of 5 groups:

1. 50% ticket subsidies for 10 weeks
2. 50% ticket subsidies for 20 weeks
3. 75% ticket subsidies for 10 weeks
4. 75% ticket subsidies for 20 weeks
5. No subsidies

Commuters will be surveyed at baseline and endline to understand the previous day's travel patterns, as well as through a midline SMS survey. We also have access to the smartcard backend to measure each trip on the formal system. The SMS survey will be used to verify which trip was taken from a set of options, to try and identify card sharing.

There will be two main empirical approaches we will follow. One is a simple reduced form diff-in-diff regression where we regress the change in outcomes on a dummy for whether the individual was in a treatment group. The core outcomes will include whether they took the BRI, the number of trips taken. Additional outcomes will include the distance of the main trip, number of trips taken using car and informal minibus, as well as other "real" outcomes such as income and hours worked. We will include control variables to increase the precision of our estimates, including respondent demographics (age, gender, education, income, number of children) and locational attributes (distance to closest BRI stop, local government area fixed effects). We will also conduct a heterogeneity analysis interacting the treatment variable with (i) gender and (ii) income.

The second empirical approach will be structural. The simplest model of transport demand we have in mind is a discrete choice model

$$U_i(\omega) = \alpha - \gamma p - \eta t + \epsilon(\omega)$$

where $i \in \{Bus, NoBus\}$ and mean utility for the no bus option is normalized to zero. Here α is an amenity from the bus, γ is the marginal utility of income (or disutility from paying fares) and η is the value of time. p is the bus fare, and $t = t^V + t^W$ is total travel time (wait time plus in-vehicle time). If ϵ is drawn from a T1EV with unit shape, then

$$P\left(Bus\right) = \frac{\exp\left(\alpha - \gamma p - \eta t\right)}{1 + \exp\left(\alpha - \gamma p - \eta t\right)}.$$

The RCT provides exogenous variation in p that allows us to estimate $dP(Bus)/dp$ (i.e. γ) using maximum likelihood.

3 Wait Time RCT

The second RCT aims to measure η from the model above, i.e. individuals' sensitivity to time.

The experiment works as follows. Our team has developed an SMS-based app which allows us to measure participants' value of time at bus stops. After being recruited at their home on weekends (which is important, so as not to select participants with low value of time willing to speak with enumerators at bus stops during peak times), participants show up to their registered bus stop and find the enumerator waiting there. The enumerator will hold up a tablet, which displays a random code which changes every minute. The participant texts this code to our SMS-service. The service will send an immediate airtime reward for checking in, as well as an offer to wait X minutes for a payment of Y Naira. To accept the offer, a participant will wait at the bus stop and then send back the new code that the enumerator's tablet displays after X minutes, allowing us to verify they had waited the requisite amount of time. If the participant does not accept the offer, they can continue with their day.

By observing the combinations of wait time and payments that are accepted and rejected, we are able to identify participants' value of time. Suppose a commuter is commuting through a bus stop. First, they decide whether to check in with us. If they check in, we will pay them an immediate amount $s_{checkin}$, but checking in incurs a slight hassle, with a wait of $t_{checkin}$. Second, to those who check in we send a wait offer, to pay them s to wait for t^W minutes. Third, the commuter decides whether to accept or reject the offer.

Let's work backwards. Conditional on checking in, she will accept the offer if

$$\alpha - \gamma(p - s) - \eta(t + t^W) + \epsilon(\omega) > \alpha - \gamma p - \eta t + \epsilon(\omega)$$
$$\Leftrightarrow s/t^W > \eta/\gamma$$

If compliance were perfect (participants always checked in when commuting) then this is the only equation we would need. But if compliance is sufficiently far from perfect then we may need to model commuters' decision to check in. They will incur the additional hassle, but after that receive the checkin payment and possible wait benefits. Our idea is to offer individuals different combinations of (s, t^W) in order to identify η/γ. We additionally will vary $s^{checkin}$. The price experiment separately identifies γ, allowing us to recover η.

These parameters can be estimated through maximum likelihood.The utility of waiting (conditional on checking in) is based on the realized wait time offer, and a noise term ν_{it} which allows individuals to make different decisions on different days,

$$D^*_{it} = \gamma s_{it} - \eta t^W_{it} + \nu^{wait}_{it}.$$

The utility of checking in is given by

$$C^*_{it} = \gamma s^{checkin}_{it} - \eta t^{checkin} + \mathbb{E}[D^*_{it} | D^*_{it} > 0] + \nu^{checkin}_{it}$$

where at the point the commuter is at the bus stop deciding whether to check in, they know ν^{wait}_{it} but not the specific offer, so this can be written

$$C^*_{it} = \gamma s^{checkin}_{it} - \eta t^{checkin} + \mathbb{E}[\gamma s_{it} - \eta t^W_{it} | \gamma s_{it} - \eta t^W_{it} > 0] + \nu^{wait}_{it} + \nu^{checkin}_{it}$$

Now, define $\omega_{it} = \nu^{wait}_{it} + \nu^{checkin}_{it}$.

We assume that the noise terms are jointly distributed according to some CDF F,

$$\begin{bmatrix} \nu^{wait}_{it} \\ \omega_{it} \end{bmatrix} \sim F(\cdot | \sigma^{wait}, \sigma^{checkin}, \rho).$$

where we assume the distribution has parameters that control the spread of the two primitive errors (σ^{wait} and $\sigma^{checkin}$), and a parameter that controls the correlation between them (ρ). Let F^x indicate the marginal CDF along error x.

Then in our data we will observe the decision to check in

$$C_{it} = \mathbb{I}\{C^*_{it} > 0\}$$

and, only if the person checks in ($C_{it} = 1$), the decision to accept

$$D_{it} = \mathbb{I}\{D^*_{it} > 0\}.$$

The likelihood is as follows

$$\mathcal{L}(\gamma, \eta, \sigma^{wait}, \sigma^{checkin}, \rho | C_{it}, D_{it}, s_{it}, s_{it}^{checkin}, t_{it}^W) = \prod_{it} \Big(1 - F^{\nu^{wait}}(-\bar{D}) - F^{\omega}(-\bar{C}) + F(-\bar{D}, -\bar{C})\Big)^{C_{it} D_{it}}$$
$$\Big(F^{\nu^{wait}}(-\bar{D}) - F(-\bar{D}, -\bar{C})\Big)^{C_{it}(1-D_{it})}$$
$$\big(F^{\omega}(-\bar{C})\big)^{1-C_{it}}$$

where participation thresholds are given by $\bar{D}(\gamma, \eta) = \gamma s_{it} - \eta t_{it}^W$ and $\bar{C}(\gamma, \eta, s^{checkin}) = \gamma s_{it}^{checkin} - \eta t^{checkin} + \mathbb{E}[\gamma s_{it} - \eta t_{it}^W | \gamma s_{it} - \eta t_{it}^W > 0]$. The log likelihood is given by

$$l(\gamma, \eta, \sigma^{wait}, \sigma^{checkin}, \rho | C_{it}, D_{it}, s_{it}^{checkin}, s_{it}, t_{it}^W) = \sum_{it} C_{it} D_{it} \log \Big(1 - F^{\nu^{wait}}(-\bar{D}) - F^{\omega}(-\bar{C}) + F(-\bar{D}, -\bar{C})\Big) +$$
$$C_{it}(1 - D_{it}) \log \Big(F^{\nu^{wait}}(-\bar{D}) - F(-\bar{D}, -\bar{C})\Big)$$
$$(1 - C_{it}) \log F^{\omega}(-\bar{C})$$

Note a few things:

- The check in offer $s_{it}^{checkin}$ is excluded from the waiting decision and so acts as an instrument for the decision to check in.
- The decision D_{it} is invariant to the scale of D_{it}^*, so these decisions will not jointly identify γ, η, and σ^{wait}. We intend to fix σ^{wait}, following common practice.
- The check in decision depends on commuters' beliefs about the utility value of the distribution of wait offers $\mathbb{E}[\gamma s_{it} - \eta t_{it}^W | \gamma s_{it} - \eta t_{it}^W > 0]$.
- This can be approximated by a simpler likelihood if compliance is very high ($C_{it} = 1$ for most observations). (It would also simplify if the errors ν_{it}^{wait} and ω_{it} were uncorrelated, but because of the mechanical correlation, that would require that ν_{it}^{wait} and $\nu_{it}^{checkin}$ be negatively correlated.)

$$l(\gamma, \eta, \sigma^{wait} | D_{it}, s_{it}, t_{it}^W) \approx \sum_{it} D_{it} \log \Big(1 - F^{\nu^{wait}}(-\bar{D})\Big) + (1 - D_{it}) \log \Big(F^{\nu^{wait}}(-\bar{D})\Big)$$

3.1 Offer distributions

Participants must use beliefs about the distribution of offers in order to compute the term $\mathbb{E}[\gamma s_{it} - \eta t_{it}^W | \gamma s_{it} - \eta t_{it}^W > 0]$. Early participants were recruited without being given detailed information about these distributions. For future participants, we will express several statistics of these distributions, such as the range of wait times $\min(t_{it}^W)$ and $\max(t_{it}^W)$ and payouts $\min(s_{it})$ and $\max(s_{it})$, and the average payout $\mathbb{E}[s_{it}]$. Although we draw from a discrete distribution, it may be difficult to optimize the likelihood over a discrete distribution, and in any case, participants are likely to hold smoother beliefs. So we plan to approximate beliefs about the distribution (s_{it}, t_{it}^W) with a smooth distribution, such as a bivariate Beta that matches the statistics presented to participants.

Our power calculations suggest that power is maximized when the fraction of offers accepted is 0.5. This is intuitive: if all offers are accepted, or all are rejected, for example, this only puts a bound on the value of time and does not allow estimation of a point estimate. While we have conducted pilots to get some information on users value of time, we may update the distribution of offers up a number of times during the RCT to aim for a fraction of accepted offers closer to 0.5.

We adjusted this distribution in the beginning of the experiment to titrate the willingness to pay, at the bus-stop level, and use other bus stops in the same "zone" (4 large geographic areas of the city) to compute users' implied value of time (we index our distributions by average value of time (in terms of Naira per minute), so if we see a fraction of offers accepted > 0.5 we will adjust to a lower Naira per minute distribution). For bus stop b we adjust this distribution based on the choices made at other similar bus stops b' (leaving out b) to avoid previous bus stop participants' choices affecting the future distribution of offers. We can conduct robustness where we compute standard errors to account for dependence. However, note that such adjustments change beliefs and thus affect the attrition correction.

3.2 Measurement

The model above incorporates the concern that individuals may select into participating based on their value of time (either across individuals, or within individuals across days). Each day a participant plays we compensate them with a small payment $s_{it}^{checkin}$ for the time to check in. This checkin amount will be drawn at the beginning of the experiment at random for each individual, to document how sensitive participation is to rewards.

We also need to measure the decision to check in (C_{it}) on days where a person is commuting. We will do this in three ways. First, in the baseline, we will ask participants which particular days in the coming week they plan to travel through the designated bus stop. Second, in the baseline, we will ask them how many days they plan to travel through the bus stop weekly in the coming weeks. Finally, in an endline survey, we will ask what days they traveled through the bus stop in the previous week. The endline question has the advantage that it will reveal actual (rather than planned) travel, but we may not be able to reach all participants. So we plan to rely more heavily in the baseline surveys, which by definition will include all of our participants. These will be subject to some planning error. If a person says they will commute $\sum_t C_{it}$ days but they check in a fewer number of days, we will add observations corresponding to those missing days with $C_{it} = 0$ (note we will not know which days were missing, so this observation cannot include time varying covariates). Alternately, if the person shows up more days than they had planned, then we will treat them as having no missing days.

In the baseline survey, we will also ask subjective value of time measures. We will assess whether attrition is correlated with these measures.

We will measure the hassle of checking in $t^{checkin}$ by measuring the additional time it takes one of our field officers to check in that is additional to the time that would normally be spent commuting. We will repeat this at a sample of bus stops. (The first time a participant checks in it may be an additional hassle to locate the enumerator, so the first check in offer is set up differently, as an onboarding task: a special higher checkin payment is given, and is not followed by a wait offer.)

Another concern is whether individuals will adjust the time they arrive at the bus stop in order to wait and receive payments. For a random subset of participants, we will give them an extra day at the beginning of the study when they receive a payment without having to wait at all. We will test for differential arrival times for this group on this day to see if this scheduling adjustment is happening.

3.3 Extensions

Assuming the implementation of this base design runs sufficiently smoothly, we will conduct extensions of this simple model:

1. Heterogeneity: allowing value of time to vary with income, either by adding in group- or income-specific coefficients (i.e. essentially re-estimating the same model for different groups) or by dividing the fare by worker income, which delivers a value of time proportional to income.

2. Individual-level Coefficients: We have included a sub-sample of 80 participants (out of the 400 we expect to play) which will play for 5 weeks (rather than 3 for everyone else). Our power calculations suggest this should be enough to uncover individual-level estimates of γ/η, which would allow us to recover the distribution non-parametrically. Otherwise we can estimate a random coefficients model a la BLP where we parameterize the distribution of these as normal and allow the parameters to vary across groups such as gender and income/education.

3. Non-Linear Value of Time: It is possible that people really dislike long waits, so we will estimate an extension which allows for a non-linear value of time. This can be incorporated by including the quadratic term $-\eta_2(t_{it}^W)^2$ in the specification of D_{it}^*

4. Distaste for uncertainty. See next section.

3.4 Model with Desired Arrival and Scheduling

Now let's enrich the model to allow commuters to have preferences over when they arrive at their destination. This section omits the compliance correction above (it assumes that commuters check in when commuting).

The commuter wishes to arrive at a given time t^{A*} and faces a cost if she instead arrives at t^A,

$$U_i(\omega) = \alpha - \gamma p - \eta t - \beta(t^A - t^{A*})^2 + \epsilon(\omega)$$

where for now we assume that cost is quadratic, and moderated by scheduling preference parameter β.

The commuter does not choose the arrival time, but can influence it by choosing a departure time t^D, and whether to accept an offer to wait an additional t^W minutes for an offer of s,

$$t^A = t^D + D_{it} \cdot t^W + t$$

where t represents trip time, which is a random variable.

The problem proceeds in two stages. Let's consider them in reverse. Once the commuter reaches the bus stop, she has already chosen t^D and sees the wait offer (s_{it}, t^W_{it}), and decides whether to accept (D_{it}) to maximize

$$\mathbb{E}[U_i(\omega)|t^D] = \alpha - \gamma\left(p - D_{it} \cdot s_{it}\right) - \eta\left(\mathbb{E}[t] + D_{it} \cdot t^W_{it}\right) - \beta\mathbb{E}\left[(t^D + D_{it} \cdot t^W_{it} + t - t^{A*})^2 | t^D\right] + \mathbb{E}\epsilon(\omega)$$

Let $D^*_{it}(t^D, s_{it}, t^W_{it})$ represent her strategy.

Previous to that, at home, the commuter chooses departure time t^D to maximize

$$\begin{aligned}\mathbb{E}U_i(\omega) =& \alpha - \gamma\left(p - \mathbb{E}\left[D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot s_{it}\right]\right) - \eta\mathbb{E}\left[t + D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W\right] \\ & - \beta\mathbb{E}\left[(t^D + D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W + t - t^{A*})^2\right] + \mathbb{E}\epsilon(\omega)\end{aligned}$$

Now consider two regimes for setting wait offers.

3.4.1 Scheduled Departures

Under scheduled departures, commuters are given a schedule of departure times $\mathbb{T}^D$, such as every $d = 20$ minutes (:00, :20, :40). Let $\tilde{t}^D$ represent the closest departure time to when a commuter wishes to go. Then the wait time is predictable,

$$t^W = \tilde{t}^D - t^D$$

In order to allow commuters to make more predictable plans, we use fixed monetary offers, $s_{it} = s_i$. At the bus stop, the commuter will wait if

$$\gamma s_i > \eta(t^D - \tilde{t}^D) + \beta\mathbb{E}\left[(\tilde{t}^D + t - t^{A*})^2\right] - \beta\mathbb{E}\left[(t^D + t - t^{A*})^2 | t^D\right]$$

Note that no further uncertainty about the offer is revealed at the bus stop. So at home, the commuter jointly chooses t^D and whether she plans to accept the offer, D_{it}. If the commuter plans to accept the offer, she will align with the scheduled departure ($t^D = \tilde{t}^D$) and expect utility

$$\alpha - \gamma(p - s_i) - \eta\mathbb{E}[t] - \beta\mathbb{E}\left[(\tilde{t}^D + t - t^{A*})^2\right] + \mathbb{E}\epsilon(\omega)$$

note this has no additional wait time.

Alternately, if she plans to reject the offer, she will expect utility

$$\alpha - \gamma p - \eta\mathbb{E}[t] - \beta\mathbb{E}\left[(t^D + t - t^{A*})^2\right] + \mathbb{E}\epsilon(\omega)$$

In this case, first order conditions yield an optimal departure time

$$t^{D^*} = t^{A*} - \mathbb{E}[t]$$

so that she plans to arrive on the right time on average. Note that quadratic loss implies that people dislike being early as much as they dislike being late; if they dislike being late more, then they would instead shade their departure time to show up early on average.

Then the commuter will accept the offer if

$$\gamma s_i > \beta\left((\tilde{t}^D - t^{A*})^2 + 2\mathbb{E}[t](\tilde{t}^D - t^{A*}) + \mathbb{E}[t]^2\right)$$

Now define the schedule disruption $\Delta := \tilde{t}^D - t^{D^*} = \tilde{t}^D - t^{A*} + \mathbb{E}[t]$, which represents the disruption from the schedule the commuter would have taken without the wait. If the person would have chosen to depart at $\tilde{t}^D$ anyway, then $\Delta = 0$. Then the above implies an optimal strategy

$$D^*_{it}(s_i, \Delta) = 1\{\gamma s_i > \beta\Delta^2\}$$

So that she will accept the scheduled departure if the value of the payment exceeds the cost of the schedule disruption.

If the commuter adjusts to the schedule, we won't measure t^{D^*}, only $t^D = \tilde{t}^D$. However, we can get at this by looking at patterns in check in times. For each individual we can look at their distribution of departure times during the random treatment $t^D_{it} \mod d$, which we anticipate will be distributed uniformly, and the difference with $\tilde{t}^D_{it} \mod d$ gives the amount of disruption.

3.4.2 Random Departures

Now consider instead the case where wait offers are realized at the bus stop. After arriving at t^D and observing a wait offer (s_{it}, t^W_{it}), the commuter's optimal strategy is

$$D^*_{it}(t^D, s_{it}, t^W_{it}) = 1\left\{\gamma s_{it} > \eta t^W_{it} + \beta\left[\mathbb{E}\left[(t^D + t^W_{it} + t - t^{A*})^2\right] - \mathbb{E}\left[(t^D + t - t^{A*})^2\right]\right]\right\}$$

or expanding, she will wait if the payment exceeds the cost of a longer trip plus the cost of any additional deviation from the desired arrival time,

$$D^*_{it}(t^D, s_{it}, t^W_{it}) = 1\left\{\gamma s_{it} > \eta t^W_{it} + \beta\left[(t^W_{it})^2 + 2t^W_{it}\left(t^D + \mathbb{E}[t] - t^{A*}\right)\right]\right\}$$

Now, when leaving home, the commuter chooses departure time t^D to maximize $\mathbb{E}U_i(\omega)$, given beliefs about the distribution of offers. If we had some idiosyncratic error to rationalize random acceptances of a given offer (s_{it}, t^W_{it}), so that $D^*_{it}(t^D, s_{it}, t^W_{it})$ was the probability of accepting, then we could generate first order conditions are given by

$$\begin{aligned} 0 =& \gamma\mathbb{E}\left[\frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot s_{it}\right] - \eta\mathbb{E}\left[\frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it}\right] \\ &- 2\beta\mathbb{E}\left[(t^D + D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it} + t - t^{A*})(1 + \frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it})\right] \end{aligned}$$

$$t^D \cdot 2\beta\mathbb{E}\left[(1 + \frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it})\right] = \mathbb{E}\left[\frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot \left(\gamma s_{it} - \eta t^W_{it}\right)\right] - 2\beta\mathbb{E}\left[(D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it} + t - t^{A*})(1 + \frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it})\right]$$

$$t^D = \frac{\frac{1}{2\beta}\mathbb{E}\left[\frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot \left(\gamma s_{it} - \eta t^W_{it}\right)\right] - \mathbb{E}\left[(D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it} + t - t^{A*})(1 + \frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it})\right]}{\mathbb{E}\left[(1 + \frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it})\right]}$$

This is complicated. Imagine $\frac{\partial D^*_{it}}{\partial t^D}(t^D, s_{it}, t^W_{it}) = 0$ so that departing earlier doesn't make a commuter more likely to accept a wait offer (we might be able to assume this is approximately zero). Then we get a simple equation,

$$t^D = t^{A*} - \mathbb{E}\left[t + D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it}\right]$$

so that the commuter departs so that on average she will arrive at their desired arrival time, anticipating the wait time offers she would accept under her belief about the distribution of offers. We observe individuals on some initial days before they are incentivized to wait. When they are incentivized to wait, this model would suggest that checkin times should move forward by an average of $\mathbb{E}\left[D^*_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it}\right]$. We can compute an analogue with the empirical acceptance distribution, $\frac{1}{N}\sum_{it}\left[D_{it}(t^D, s_{it}, t^W_{it}) \cdot t^W_{it}\right]$.

4 Further Analysis

The first part of the project will provide reduced form evidence on the impact of the BRI on the informal network, and estimate a model of demand and supply for transit using the RCTs to measure the total impact of the system on consumer surplus and its distributional effects accounting for impacts on drivers. A key aspect of this approach is it will capture indirect effects of formal transit provision (since users of the informal system may be affected if prices and supply responds) and impacts on emissions (since changes in supply will impact emissions if danfo are particularly polluting, or if users near the margin of driving are induced to use the new formal system).

The second project will use the same RCTs to estimate a model of a decentralized transport network and characterize optimal policies to regulate prices and entry in the industry. This is an important question because the solutions of formal mass rapid transit like Bus Rapid Transit and Light Rail are very expensive and not a medium-run solution for improving public transit in Sub-Saharan Africa. The backbone is a dynamic queueing model which captures the various external effects present in the transit market (increasing returns to scale or "wait time" externality; business stealing; network effects through commuters combining multiple routes via transfers). The researchers will use the model to characterize how governments should regulate prices and entry to maximize societal welfare, inclusive of both consumer surplus and emissions.