Revealed and Stated Preference Modeling of Transportation Demand:

A Control Function Approach

May 2022

by

Felix L. Friedt
Macalester College
ffriedt@macalester.edu

Wesley W. Wilson
University of Oregon
wwilson@uoregon.edu

Abstract

The demand for freight transportation plays a central role in evaluating infrastructure investments and the effects of pricing decisions. In this study, we use survey data of agricultural shippers to estimate transportation demand decisions using a choice model where shippers choose the mode and the destination which define the shipment. In agricultural markets, there are an enormous number of destinations and shippers face a different set of modal options. In our data, shippers reveal the relevant choice set (alternative mode/destination pairs that they consider) and the decision made. The decision rests on a set of attributes, and as is commonly observed with revealed data for which the range of attributes can be limited. We follow recent research wherein revealed data are supplemented with a set of stated preference data to estimate the choice model. In our case, the stated preference data are based on the revealed data and make the stated preference data endogenous. We control for the endogeneity of stated preference data using a control function approach, and apply the model to agricultural shippers in the Ohio River Basin. We find that prices received for the product shipped, the transport rate, the time in transit, and reliability of the mode each have statistically important effects on decisions. The results are then used to find that modal demands are relatively inelastic with respect to each of the attributes as well as modes and point to a considerable number of captive shippers.

Keywords: Transport Demand, Revealed Preference, Stated Preference, Choice Model, Captive Shipper

---

The authors gratefully acknowledge, comments and data support received from Eva Opitz, William Davis, Brian Shepard, and June Gibbs of CDM Smith. The research uses survey data under OMB Control Number 0710-0022. Earlier drafts of this study were presented at the American Economic Association Annual Meetings in 2022 and the University of Saskatchewan Department of Agricultural and Resource Economics Research Seminar; we thank session attendees, and in particular Alejandro Molnar and James Nolan, for their valuable feedback.

0. Statements and Declarations

On behalf of all authors, the corresponding author states that there is no conflict of interest.

1. Introduction

There is considerable interest in transport demand as policymakers attempt to find solutions to an aging infrastructure, congestion, environmental degradation, changing regulations etc. Solutions include capital investments, congestion fees, and carbon taxes. Such issues and policies directly affect the transport markets, and the prices shippers pay for transportation. As stated by the National Research Council (NRC 2005), "Price responsiveness is so important to estimating the benefits of [infrastructure] improvements that informed judgments about the merit of such improvements cannot be made without careful study of these demand and supply elasticities."

This interest in and importance of transportation has resulted in a number of demand studies using different techniques and data. Recent advances in this field of study include, for example, research by Kalahasthi et al. (2022), Keya et al. (2019), or Khakdaman et al. (2020). While all of these papers consider and evaluate the demand for freight transport, each takes a different approach. Kalahasthi et al. (2022), for example, develop a Freight Origin-Destination Synthesis with Mode Choice (FODS-MC). The key innovation is the addition of mode choice, which allows the authors to jointly estimate the trip distribution, mode choice, and empty trips in truck and rail markets. Keya et al. (2019) also study freight transport demand but do so via a Copula-based random regret minimization framework. The innovation of in this study is the joint determination of model choice and shipment size and offers substantial improvement in model fit over more simplistic frameworks. The authors show a strong connection between mode choice and shipment size, a finding that is echoed in our research. Finally, Khakdaman et al. (2020) examine whether shippers are willing to delegate the selection of modal choice. The findings point to a recent paradigm shift that has increased the willingness of shippers to concede control over modal choice to logistics service providers. While this finding is intuitive for shippers seeking global market access and therefore entering international transport markets (i.e. Fortune 500 companies), mode choices are likely much more limited for smaller shippers targeting domestic markets. In fact, our study of agricultural shippers indicates that many firms have very few alternatives (mode and/or destination) and that shutting down in response to worsening transport conditions is a viable option for many shippers.

In this study, we focus on transportation demand in agricultural markets. In these markets, there have been a number of studies, using an array of different techniques and data. Following the classic works of Oum (1979) and Friedlaender and Spady (1980), Wilson (1984), Dybing (2002) and others use demand functions that are derived from a transportation cost function using data aggregated over shipments. Other studies use aggregated data and a spatial equilibrium framework. These include studies by Yu and Fuller (2005; 2006) who estimate the demand for agricultural shipments by barge in the Upper Mississippi waterways, while Babcock and Gayle (2014) estimate rail demand between east and west region in the United States, each using a spatial equilibrium approach. In other studies, Miklius et al. (1976) uses a logit model on aggregated shares for cherry and apple shipments in the Pacific Northwest. Fitzsimmons (1981) use aggregate data on rail shipments to estimate the demand for grain and soybeans. Wilson et al. (1988) use aggregated data for shipments of rail and truck from North Dakota to Minnesota destinations. Miljkovic et al. (2000) estimate the demand for rail and barge shipments from Illinois to the Gulf States in a system of demand and supply equations. Inaba and Wallace (1988) and Henrickson (2011) use a model of spatial competition between elevators. While Inaba and Wallace (1988) use a mixed continuous/discrete model, Henrickson (2011) uses a spatial auto-correlation model. Finally, there has been a number of studies by Train and Wilson (2004; 2006a; 2008; and 2019) that estimate demands using disaggregated survey data that applies to shippers in the Pacific Northwest and the Upper Mississippi River Basin.

In this study, we use disaggregated survey data, patterned after the Train and Wilson studies, for the Ohio River Basin. In this region, as in most agricultural markets, there are an enormous number of locations (origins and destinations) to which commodities flow. These flows are typically done by truck, rail, barge, or some combination thereof. The previous research uses aggregated data or disaggregated data in modeling demands.[1] In studies that use aggregated data, the aggregation occurs over a region e.g., grain movements from ND to MN or IL to the gulf states, etc. or across shipments. In both cases, there is considerable information that can be lost in the aggregation. Oum (1979; 1989) and Winston (1983), for example, point to the lack of behavioral aspects and/or biases that arise from the functional forms. Furthermore, often the aggregation involves an artificial market boundary e.g., county, state, etc. and requires the use of a "representative" shipper, which masks the heterogeneity of shippers. In studies that use disaggregated data, the focus is typically on mode choice across different shippers. Disaggregate models overcome these issues by modeling the choice that shippers make from an array of options available to them. Such studies can be complicated by the lack of sufficient variation in the attributes affecting the revealed choice, and researchers have introduced stated preference modeling as a way to introduce greater variation.

In our case, we use the survey data to estimate transportation demands, using a combination of revealed and stated preference data following Train and Wilson (2008) and Guevara and Hess (2019).[2] Train and Wilson (2004; 2006a; 2008; and 2019) provide analyses of the same type of data and use a full-information maximum likelihood estimator to estimate transport demands in the Upper Mississippi River Valley and the Columbia-Snake River Valley. In our case, the data are quite limited and we were not able to estimate with the full-information maximum likelihood estimator. Instead, we use a control function

[1] Winston (1983), Oum (1979; 1989), Small and Winston (1998) Clark et al. (2005) and others describe alternative approaches to estimating transport demands.

[2] Train and Wilson (2004; 2006a; 2008; and 2019) provide analyses of the same type of data and use similar techniques for the Upper Mississippi River Valley and the Columbia-Snake River Valley. Unlike these studies, our study uses a control-function estimator developed by Guevara and Hess (2019).

approach developed by Guevara and Hess (2019) which they applied to the Train-Wilson (2008) data for the Columbia-Snake River Basin.

Theoretically, there are an enormous number of outlets for agricultural products and there are a number of modal options available to shippers, which makes the choice set very large. Yet, for a specific shipper only a handful of options may be relevant. In our case, the choice set is solicited from the shipper along with the choice. The choice is made on a utility basis which is framed in terms of the price received by the shipper for the product transported, the transportation rate, time in transit, the reliability of the mode, and shipper characteristics.

As has become recent practice, stated preference data are combined with the revealed data. Stated preference data can be used to insure sufficient variability in the attributes, and construction of the stated preference data from the revealed choice helps to provide realism to the stated preference experiments.[3] But, as pointed out by Train and Wilson (2008), construction of the stated preference data from the revealed choice, makes the stated preference data endogenous. Train and Wilson (2008) developed a full information maximum likelihood estimator for the use of stated preference data constructed from the revealed (SP-off-RP) data which is implemented with simulation. As noted by Guevara and Hess (2019) (G-H), the Train-Wilson (T-W) approach is difficult to apply and requires large samples. To overcome this issue, G-H develop a limited information maximum likelihood estimator to address the problems of SP-off-RP data, which does not need simulation. Their estimator uses a control-function approach to correct for the endogeneity and can be estimated with conventional software. This is particularly attractive in the present case as the sample size is somewhat limited.

[3] Rose et al. (2005), Hensher and Greene (2003), Hensher (2004; 2006) and Caussade et al. (2005) use a pivot design where in stated preference questions are constructed from the revealed decision. Fowkes and Shinghal (2002), Bergantino and Bolis (2006), Fowkes (2004) use an "adaptive" approach wherein the stated preference data are constructed from the revealed-preference data by making the revealed choice worse.

We describe the empirical model in Section 2. Our analysis applies to agricultural shippers in the Ohio River Basin. The data were obtained through a survey of shippers located in the region, where 190 responses were obtained and used to estimate the models. The survey, variable definitions and summary statistics are provided in Section 3. Section 4 provides the results, which include estimates of the coefficients using only the revealed preference data, only the stated preference data, and the combined data using the G-H approach.

The descriptive results themselves are of direct interest. There are considerable differences in terms of modal access which suggest that only about 33 percent of the grain elevators have access to rail, and only about 11 percent have access to barge. This means that truck is the primary mode in the analysis, and in terms of modal decisions dominate barge and rail. The results also point to an enormous number of locations to which shippers ship. Indeed, in terms of shipments made, there are about 185 shippers that provided the information, and there were 104 unique locations to which they ship. Among these locations processing plants, river terminals, and other rail/elevator terminals are the predominant shipment destinations. The elevators themselves have been in business 63 years on average with relatively few that have been in business 10 years or less, which suggests the elevators have considerable logistic experience and practices. There is a wide variation of shippers in terms of tons shipped and storage capacity, but most are relatively small.

The primary econometric findings point to the statistically significant impact of the observed shipment attributes. Demand for transport rises with price received for the sold commodity at the destination. As expected, the transportation costs (rate) have a significant negative impact on the probability of choosing a given mode-destination pair. Similarly, an increase in the time in transit and distance shipped or worsening of the reliability (on-time delivery) lead to a reduction in the demand for transport.

While these coefficient estimates are statistically important and carry the expected signs, in all cases, the calculated elasticities are quite small. Overall, changes in reliability prompt larger responses from shippers than changes in rate or time-in-transit. Shippers choosing trucks are least responsive to changes in freight rates but most affected by changes in reliability. Changes in time-in-transit trigger the largest responses from shippers choosing barge followed by those choosing rail. It appears grain shippers in the Ohio River Basin tend to be captive; unable to switch mode and/or destination in response to worsening transport market conditions.

Because the demand for freight transportation plays a central role in evaluating infrastructure investments and the effects of pricing decisions, it is critical to understand how shippers will respond to large scale investments such as those proposed through the Build Back Better plan. Our study produces new estimates of the elasticities of demand for transport in U.S. agricultural markets with respect to multiple shipment attributes. The fact that we find evidence of inelastic demand for transport points to captive shippers and has policy relevance. In this environment, an improvement of existing infrastructure that lowers the cost of transport, for example, has benefits, but these are limited given the inelastic response of shippers. Investments that diversify transport markets and improve access to multiple modes may be more effective because it intensifies competition in transport markets and reduces the monopoly position of transport firms holding shippers' captive.

2. Choice Modeling with a Control Function

In this section, we provide an overview of the Guevara and Hess (2019) approach to estimating combined stated and revealed preference data where the stated preference data (sp) are constructed from experiments based on the revealed data (rp). As noted earlier, RP data are sometimes limited in the range of the attributes. In response, researchers began to use stated preference data. While this approach allowed them to introduce the desired degree of attribute variation, it may be subject to a lack of realism to the respondents. The SP-off-RP design constructs the SP data from the revealed choice which grounds the stated preference response to the revealed choice, providing realism to the experiment. However, such an approach can result in a dependence between the SP attributes and the unobserved factors which runs contrary to the independence assumption for standard estimation procedures (Train and Wilson (2008)).

In Train and Wilson (2008), Guevara and Hess (2019), and the present analysis, shippers are asked to identify the options they have and which they chose (the RP data). They are then asked which option they would choose in the RP setting if the rate, time, or reliability of the option they actually chose were changed. These questions have two features that need to be addressed in the estimation. First, when answering the SP-off-RP questions, the shipper is choosing among options in the RP setting. This implies that the attributes of the options in the RP setting, including, importantly, the attributes that are not observed by the researcher, affect the shipper's answer to the SP-off-RP questions. Stated in econometric terms, the unobserved factors associated with each option in the RP setting can be expected to enter the shipper's evaluation of these options when answering the SP-off-RP questions. Second, the SP-off-RP questions ask the respondent about a change in the rate, time, or reliability of the option that was chosen in the RP setting. In econometric terms: The SP-off-RP questions are conditional on the outcome of the RP choice. This conditionality implies that the distribution of unobserved attributes that enter the shipper's responses to the SP-off-RP responses is not the unconditional distribution, as in standard choice models, but rather the distribution conditional on the shippers' RP choice.

The econometric method incorporates both of these implications, building upon the earlier work reported in Train and Wilson (2004; 2007; and 2019). The unobserved factors in the RP setting enter the model of the shipper's response to the SP-off-RP questions, and the probability of each possible response is derived based on the distribution of these unobserved factors, conditional on the shipper's choice in the RP setting. The model is specified below and takes a two-stage control function approach to address this issue. In the first stage, the endogenous attributes of the SP-off-RP alternatives are regressed on the observed RP attributes. In other words, the RP attributes serve as instrumental variables for the SP attributes. In the second stage, the residuals of the first-stage regression are integrated in the SP-off-RP fixed coefficients choice model to control for endogeneity.

With fixed coefficients, the shipper's choice in the RP setting is a standard logit model. The shipper faces J alternatives for its last shipment, which are the alternatives that the shipper reports are available. The utility of each alternative depends on observed variables, namely, rate, transit time, and reliability, as well as unobserved factors.[4] The observed variables are denoted x_j for alternative j (with the subscript for the shipper omitted for simplicity), and the unobserved random factors are denoted collectively ε_j as for alternative j. Utility of alternative j is denoted $U_j=\beta x_j+\varepsilon_j$. Under the assumption that each ε_j is distributed iid extreme value, the probability that the shipper chooses alternative i is the logit formula:[5]

$$P_i = \frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}}$$

The researcher presents the shipper with a series of SP-off-RP questions that are constructed on the basis of the shipper's RP choice. The notation used is more general than necessary for our particular SP-off-RP questions, to facilitate the use of the method in other settings that might use different types of SP-off-RP questions. (For example, our questions ask the shipper about a change that makes the option they chose worse; an alternative would be to ask the shipper about a change that improves an option that they did not choose.) The researcher asks T SP-off-RP questions, with attributes $\tilde{x}_{jt}^{i}$ for alternative j in question t

[4] The model is framed in a utility context although the term profit maximization can be employed so long as there are no agency issues i.e., the shipper makes decisions consistent with the firm's objective of maximizing profit.

[5] This formula can be interpreted as in the following example: Suppose the shipper faces two alternatives with the observed portion of utility being βx1=3 for the first alternative and βx2=4 for the second alternative. Even though the observed portion of utility is lower for the first alternative, the shipper might still choose the first alternative because of unobserved factors. The formula states that the probability that the shipper chooses the first alternative is exp(3)/(exp(3)+exp(4))=0.27 and the probability that the shipper chooses the second alternative is exp(4)/(exp(3)+exp(4))=0.83.

based on alternative *i* having been chosen in the RP setting. For our questions, $x\tilde{}^{i}_{it} \neq x_i$ for the alternative that was chosen in the RP setting, while $x\tilde{}^{i}_{jt} = x_j \; \forall \; j \neq i$ for the non-chosen alternatives; however, more general specifications of $x\tilde{}^{i}_{jt}$ are possible. The shipper is asked to choose among the alternatives in response to each SP-off-RP question.

Herein lies the key issue with the SP-off-RP data. By design of the SP experiments some of the inherent preferences of the shippers, again denoted ε^{i}_{j} and unobserved by the researcher, transfer from the RP to the SP experiment. In addition, the shipper's choice in the SP-off-RP setting can be affected by random factors, denoted by η_j, that did not arise in the RP setting. Examples may include the inattention by the agent to the task, pure randomness in the agent's responses, or other quixotic aspects of the SP choices. Following Guevara and Hess, we model the shipper’s utility from alternative *j* in SP-off-RP question *t* as:

$$W_{jt} = \beta x\tilde{}^{i}_{jt} + \varepsilon^{i}_{j} + \eta_j.$$

That is, the shipper evaluates each shipping alternative using the same utility coefficients and with the same unobserved attributes as in the RP setting, with the addition of a new random error that reflects quixotic aspects of the shippers’ responses to the SP-off-RP questions.

Because of the transfer of ε^{i}_{j} from the RP to the SP-off-RP experiment, the SP-off-RP attributes, which are conditioned on the RP attributes, are endogenous. Guevara and Hess, however, point to the fact that (1) by construction, the RP attributes are correlated with the SP-off-RP attributes, and (2) by assumption, the RP attributes are independent of the RP random error component. Accordingly, Guevara and Hess propose a control function approach wherein RP attributes x_j (i.e., rate, time, and reliability) serve as instruments for SP-off-RP attributes $x\tilde{}^{i}_{jt}$ in a set of first-stage regressions that take the form:

$$x\tilde{}^{i}_{jt} = \alpha + \lambda x_j + \delta_{jt}.$$

Here, α represents a vector of attribute-specific intercepts and λ represents a coefficient matrix that captures the constructed correlations between RP and SP-off-RP attributes. Based on these regressions, the researcher predicts the attribute-specific residuals represented by vector δ_{jt}. These residuals capture the component of the SP-off-RP attributes that was correlated with the unobserved RP factors transferred to the SP-off-RP experiment.

In the second stage, these first-stage residuals are added to the SP-off-RP choice model to control for the endogeneity of the SP-off-RP attributes. The resulting systematic part of utility takes the following form:

$$V_{jt} = \alpha_j + \beta \tilde{x}^{i}_{jt} + \theta\delta_{jt},$$

where α_j gives a vector of alternative specific intercepts (including the option to shut down in the SP-off-RP experiment), and β and θ represent coefficient vectors to be estimated. The latter coefficients on the first-stage residuals serve as a direct test of the presence of endogeneity in the SP-off-RP attributes. A rejection of the null hypothesis (i.e., $\theta = 0$) provides evidence of the transfer of unobserved factors from the RP setting to the SP-off-RP experiment.

On a final note, Guevara and Hess point out that because of the integration of the first-stage residuals, the second stage fixed coefficient logit estimation includes estimated regressors. Consequently, the standard errors on the coefficient estimates of interest (i.e., θ) cannot be derived directly. Instead, the delta-method or a nonparametric approach, such as bootstrap, must be applied.

While the limited information maximum likelihood approach proposed by Guevara and Hess is less efficient than the original method advanced by Train and Wilson, the two-stage control function procedure produces more consistent estimates owing to the fact that it relies on fewer distributional assumptions. Moreover, the Guevara and Hess methodology does not require simulation and can be applied using conventional methods. These features make the Guevara and Hess approach an attractive and reliable alternative to the Train and Wilson approach.

3. Survey of Agricultural Shippers

The data were collected by a survey of 1150 shippers (grain elevators) in seven states including Illinois, Indiana, Kentucky, Michigan, Ohio, Pennsylvania, and Tennessee[6] in the fall of 2020. A total of 190 shippers responded to the survey, yielding a response rate of about 16.5 percent.[7] The origins are located throughout the region (Figure 1) and ship to a wide variety of destinations (Figure 2) with most in the study region. On average, the elevators have been in operation for over 60 years, and only five (of 185 responses) have been in business less than 10 years. Generally, the elevators are owned by firms which operate relatively few elevators. Indeed, 73 of 175 respondents represent single elevator operations, and 126 elevators are operated by firms that operate five or fewer elevators. There is also a significant range of reported elevator sizes. Most are relatively small where more than half of the respondents ship less than 100,000 tons and more than half have storage capacities less than 50,000 tons. However, there also a number that ship more than 1 million tons a year and 33 of 169 elevators have storage capacities of more than 100,000 tons.

[6] The survey was conducted by CDM Smith, Inc.
[7] Shippers were first invited to participate in the survey with a post card that included a weblink. Non-respondence received a second invitation with a mail version of the questionnaire, which was followed by postcard reminders to non-respondents. The remaining non-respondents were sent a second and final mailing of the mail version of the survey instrument.

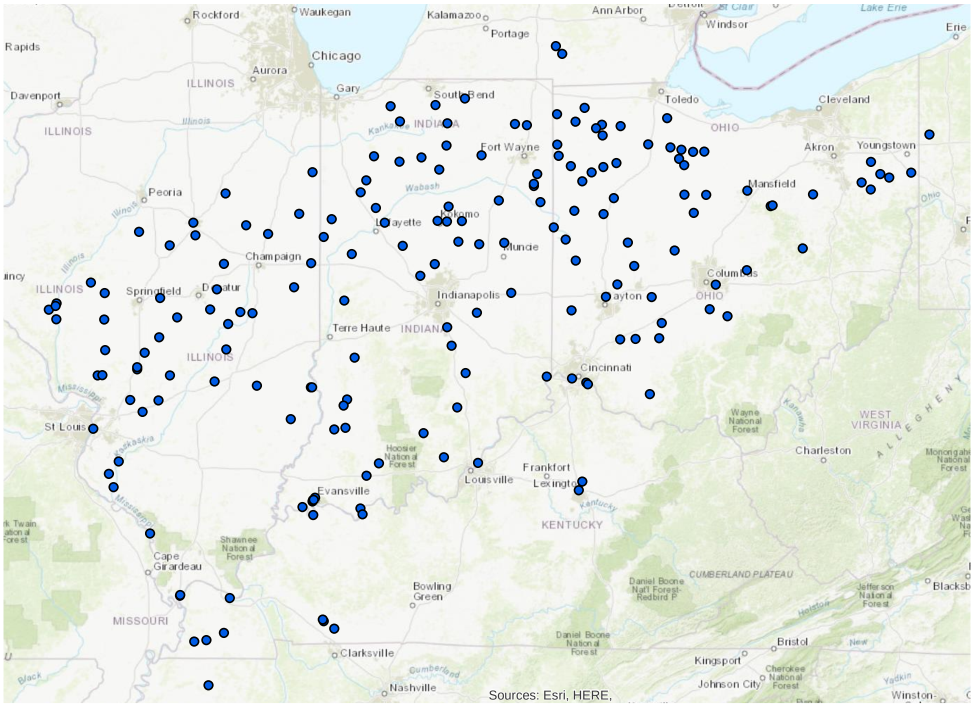

Figure 1: Shipper Locations

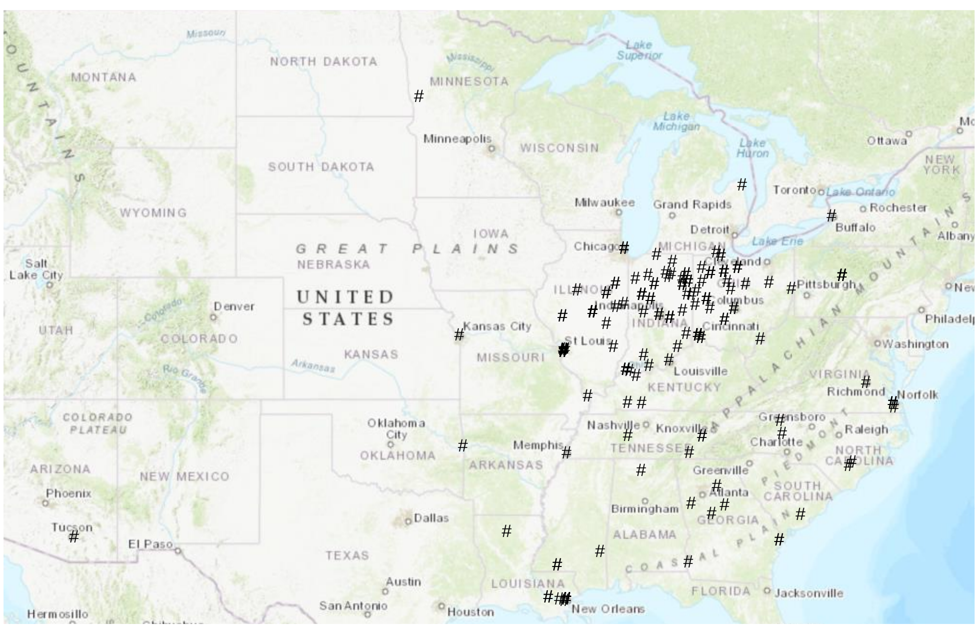

Figure 2: Shipment Destinations

Table 1 contains a summary of the modal choices made. All three modes are represented, but as noted earlier, truck shipments dominate with 135 of 186 shipments (73 percent), followed by rail with 28 of 186 shipments (15 percent), and by barge with 13 of 186 shipments (7 percent). In the questionnaire, shippers were asked of alternatives, and up to three alternatives were solicited, where again truck account for most.

Table 1: Modal Choices and Alternatives

Mode	Chosen	Percent	Alternative 1	Percent	Alternative 2	Percent	Alternative 3	Percent
Barge	13	7.0	3	2.2	3	5.8	0	0.0
Rail	28	15.1	12	8.9	0	0.0	1	3.0
Truck	135	72.6	109	80.7	49	94.2	32	97.0
Rail-Barge	2	1.1	2	1.5	0	0.0	0	0.0
Truck-Barge	5	2.7	1	0.7	0	0.0	0	0.0
Truck-Rail	3	1.6	8	5.9	0	0.0	0	0.0
Total	186	100	135	100	52	100	33	100.0

Note: For all options, some of the respondents did not identify the mode and these were excluded.

Figure 2 above summarizes the destinations of the revealed shipments made, and Table 2 below contains a summary of the destination types. While there are destinations in a wide array of states, most tend to be in the eastern portion of the study area. There are 188 shipments. The primary states include OH (39), IN (44), IL (35), and LA (20) totaling 138 of 188 shipments.[8] The remaining 50 shipments are destined in 11 other states primarily in the east. The shippers were asked not only the location to which they ship, but also the type of location. Processing plants dominate with 87 of 190 responses (46 percent), followed by river terminals with 44 of 190 responses (23 percent), and elevators (including another elevator, a rail terminal, or and export terminal), and others or not specified. The shippers also provided the type of destination for the alternatives if the original shipment could not be made. Again, these are dominated by processing plants.

[8] The number in () is the frequency observed for each state.

Table 2: Destination Alternatives

Destination Type	Chosen	Percent	Alternative 1	Percent	Alternative 2	Percent	Alternative 3	Percent
River Terminal	44	23.2	31	22.1	16	30.8	6	18.2
Another Elevator	16	8.4	13	9.3	8	15.4	4	12.1
Railroad Terminal	16	8.4	17	12.1	4	7.7	7	21.2
Processing/Ethanol Plant	87	45.8	61	43.6	24	46.2	15	45.5
Other & Not Specified	6	3.2	17	12.1	0	0.0	0	0.0
Export Terminal	11	5.8	0	0.0	0	0.0	1	3.0
Feed lot	10	5.3	1	0.7	0	0.0	0	0.0
Total	**190**	**100.0**	**140**	**100.0**	**52**	**100.0**	**33**	**100.0**

Shippers were asked to provide information on the last shipment made and to provide the same information on the next best alternative if that shipment could not be made (repeated three times). The information provided include the price received at the destination (price/ton), the transportation rate (rate/ton), the time in transit (hours) and reliability (the percentage of times that shipments like the one made arrive on time). Table 3 summarizes the results by mode and Table 4 summarizes the result by the options identified by the shipper. The rates per ton vary across the three modes, but they all ship different distances. On a per ton-mile basis, barge rates are the lowest, followed by rail and then by truck. Likewise, shipment speeds (miles per hour) are lowest for barge, followed by rail and then by truck. Finally, there are considerable differences in the reliability measure (shippers were asked for shipments like this one, what percent of the time do you expect them to arrive on time?) with truck and barge at 87 and 86 percent and rail much lower at 55 percent.

Table 3: Shipment Attributes-Descriptive Statistics by Mode

Mode	Price/Ton ($)	Rate/Ton ($)	Time (Hours)	Reliability (Percent)	Distance (Miles)	Rate/Ton-mile ($)	N
Barge	219	15.4	312	87	1,023	0.015	67
Rail	174	29.9	112	55	545	0.055	134
Truck	219	8.7	6	86	65	0.134	1,169
Rail-Barge	183	35.4	408	90	886	0.040	4
Truck-Barge	297	19.6	362	74	582	0.034	13
Truck-Rail	211	30.1	99	85	336	0.090	20
Overall	**215**	**11.6**	**37**	**83**	**167**	0.069	**1,407**

The summary statistics by option are presented in Table 4 and compares the shippers' chosen alternative with the alternative options that could have been chosen. As expected, the chosen option has higher averages for price received and reliability. The rates per ton and the transit times, measured in hours, are affected by distance. On a per-mile basis, the rate per ton-mile is lowest for the chosen option (4.8 cents per ton-mile) and larger for the other options (8.1, 10.4, and 10.4 for options 1, 2, and 3, respectively). Miles per hour is about the same for the chosen option and first alternative.

Table 4: Shipment Attributes by Option

Option	Price Received/Ton ($)	Rate/Ton ($)	Time (Hours)	Reliability (Percent)	Distance (Miles)	Average Rate/Ton-mile
Chosen Option	217.1	12.6	56.7	88	264	0.049
Alternative 1	215.2	9.1	24.6	85	113	0.081
Alternative 2	212.9	12.8	12.1	87	122	0.104
Alternative 3	208.3	12.5	13.4	83	120	0.104

The respondents were asked a series of SP-off-RP questions. In each case, the SP question was posed as whether the shipper would stay with their original choice or switch to another alternative if the attribute of the original choice was changed by X percent, where X was randomly selected from 10, 20, 30, 40, 50 and 60 percent. In all cases, the revealed option was made worse. That is, the rate and time in transit were increased, while reliability decreased.

The SP responses are summarized in Tables 5, 6, and 7 for rates, time in transit and reliability. In each case, the shipper can respond with switch to alternative, stay with original choice, or shutdown. For rates, there are a total of 181 responses. At 10 percent values of rate changes, 76 percent of responders indicated they would not switch to the alternative. As the rate change increased, this proportion fell. For large rate increases, 20 percent (4 of 20 respondents) reported they would still not switch. If they would switch, there were two possible alternatives to choose from. First, they could switch to their next best mode/destination. At various rate changes, there were a total of 64 such switches. Second, they could switch to shut down. Shut-down is and has been a major factor in all of the previous surveys conducted using this framework. In this sample, 21 of 181 (12 percent) reported that they would shut down at the rate increase prompt. As expected, switching to an alternative tends to increase with the magnitude of the rate change.

Transit times are the total of the setup and waiting times and the time once loaded to reach the final destination. There were 178 responses. If transit time rises, shippers report that a total of 106 shipments (60 percent) would not change regardless of the time change. As with rates, switching becomes more likely with progressively higher changes in transit times.

Table 7 presents the information with respect to reliability (where reliability was decreased in the SP experiment). There were a total of 178 responses. The same general patterns as with rate and time are indicated (as expected). For decreases in reliability, the switch rate generally increases with the percentage change in reliability, but unlike rates and time, there is considerable uniformity in the switch rate. Generally, Tables 5, 6, and 7 follow expectations. Further, shippers appear to be somewhat more responsive to rates than to time and reliability, particularly for large rate changes.

Table 5: Shipment Stated Preference – Rate Responses

Percent Rate Change	No Switch	Switch	Shut-Down	Total	Percent No Switch	Percent Switch	Percent Shut-Down
10	25	6	2	33	76	18	6
20	22	6	4	32	69	19	13
30	16	11	2	29	55	38	7
40	18	15	5	38	47	39	13
50	11	14	4	29	38	48	14
60	4	12	4	20	20	60	20
Total	**96**	**64**	**21**	**181**	**53**	**35**	**12**

Table 6: Shipment Stated Preference – Time Responses

Percent Time Change	No Switch	Switch	Shut-Down	Total	Percent No Switch	Percent Switch	Percent Shut-Down
10	24	4	3	31	77	13	10
20	18	6	3	27	67	22	11
30	18	7	1	26	69	27	4
40	14	13	3	30	47	43	10
50	13	8	9	30	43	27	30
60	19	13	2	34	56	38	6
Total	**106**	**51**	**21**	**178**	**60**	**29**	**12**

Table 7: Shipment Stated Preference – Reliability Responses

Percent Reliability Change	**No Switch**	**Switch**	**Shut-Down**	**Total**	**Percent No Switch**	**Percent Switch**	**Percent Shut-Down**
10	16	7	1	24	67	29	4
20	18	16	3	37	49	43	8
30	23	12	3	38	61	32	8
40	15	9	2	26	58	35	8
50	9	4	3	16	56	25	19
60	22	11	4	37	59	30	11
Total	**103**	**59**	**16**	**178**	**58**	**33**	**9**

4. Empirical Results

We use these survey responses, both RP and SP-off-RP, to estimate a model of transport demand following the G-H approach. Table 8 contains the estimation results. There are three columns. Column

(1) contains the coefficient estimates based on RP data alone; Column (2) contains the estimates on SP-off-RP data alone; and Column (3) contains the estimates based on both RP and SP-off-RP data, and accounts for the endogeneity of the stated-preference prompts following Guevara and Hess (2019). The model includes controls for mode-specific alternatives, rate (in dollars per ton), shipment times (in hours), reliability (percent on-time), price for the commodity at the destination (in dollars per ton) and distance (in miles). It also includes a shipper measure of rail car siding capacity.

In the RP results, the estimated coefficients of price, rate, time, reliability, and distance all take the expected signs, and the coefficients on time, reliability, and distance are statistically different from zero at conventional levels. Notably, the estimated coefficients on price of the commodity at the destination, and on mode specific intercepts and interactions are each not statistically different from zero at conventional levels. As is often the case, this may be the result of a lack in variation in observed destination prices revealed by shippers.

Column (2) gives the estimated parameters of a fixed-coefficients logit model estimated on the stated preference data alone, without adjustment for the potential endogeneity. A possible response to the SP-off-RP questions was for the shipper to choose to shut down. In addition to the previous independent variables and mode-specific (rail and barge) intercepts, there is a control for the shut-down alternative and its interaction with the shipper's storage capacity. Again, the estimated coefficients on price, rate, time, reliability, and distance take the expected sign, and with the exception of the coefficient estimate on reliability, are generally in line with the estimates based on the *RP* data shown in column (1). The notable change in the parameter estimate on reliability may be indicative of the underlying endogeneity in the *SP-off-RP* data. The coefficients on price, rate, time, and distance are statistically different from zero at the conventional levels. The parameter estimates on the mode-specific alternatives and the shut-down option are also statistically significant and indicate that, all else equal, rail and barge are the preferred modes of transport and the likelihood of choosing to shut down declines with shipper's storage capacity.

Column (3) gives the estimated parameters of the primary fixed-coefficients logit model estimated on the revealed preference data combined with the responses to the *SP-off-RP* questions. The estimation accounts for the endogeneity of the stated-preference prompts on the revealed preference choices using the Guevara and Hess approach.

In this case, the coefficients on price, rate, time, reliability, and distance are all of the correct sign and statistically significant at conventional levels as are the modal dummies and shutdown. Adjusting for the endogeneity, the parameter estimate on reliability once again rises and approximates the estimate found based on the *RP* data. The average utility of shutting down must be interpreted in combination with the interaction term of a shipper's storage capacity. The parameter estimates suggest that all else equal, shutting down is an option for shippers with the smallest storage capacity, while shippers with the largest elevator capacity would experience disutility from shutting down. The inclusion of the shut-down option in the model constitutes an important addition. In particular, numerous shippers stated that they had no shipping alternatives other than the one they used. For these shippers, their only alternative in the face of increasing rates or time was to shut down. Even shippers who had shipping alternatives might choose to shut down in response to potential changes in rates, time, and reliability for their chosen shipment, rather than switch to their next-best shipping alternative. In fact, many shippers responded in this way to the hypothetical changes in rates, times, and reliability. The model explicitly accounted for these responses. As the estimates indicate, the shut-down option is considered onerous for the largest of shippers, and the threshold for deciding to shut down varies considerably with a shipper's storage capacity.

Table 8: Coefficient Estimates

VARIABLES	(1) RP only	(2) SP-off-RP w/o adjustment	(3) SP-off-RP w. adjustment
Rate, in dollars per ton	-0.0023	-0.0472***	-0.0310***
	(0.0250)	(0.0131)	(0.0133)
Time, in hours	-0.0104**	-0.0045**	-0.0058**
	(0.0045)	(0.0018)	(0.0027)
Reliability	3.6612**	0.6574	2.5628***
	(1.6241)	(0.4046)	(0.6808)
Price at destination, in dollars per ton	0.0373	0.0029*	0.0030*
	(0.0252)	(0.0017)	(0.0017)
Distance, in miles	0.0044**	0.0023***	0.0024***
	(0.0018)	(0.0008)	(0.0009)
Rail constant	0.2848	1.9867***	2.1205***
	(1.5335)	(0.6569)	(0.7762)
Rail constant*log(1+rail load capacity)	0.4488	-0.1786	-0.1100
	(0.4544)	(0.1831)	(0.2002)
Barge constant	1.8969	1.5803*	1.9409**
	(2.7407)	(0.8504)	(0.8979)
Barge constant*log(1+rail load capacity)	0.1477	-0.1402	-0.1180
	(0.6933)	(0.1982)	(0.2053)
Shut-down constant		3.3791**	5.4965***
		(1.4367)	(1.5305)
Shut-down constant*log(1+storage capacity)		-0.3759***	-0.4020***
		(0.1265)	(0.1218)
First-stage rate residual			-0.0304
			(0.0412)
First-stage time residual			0.0018
			(0.0066)
First-stage reliability residual			-2.9684***
			(0.9374)
Observations	318	1,470	1,788
Log-likelihood	-88.47	-414.1	-504.1

The coefficient estimates shown in Column (3) of Table 8 allow for the calculation of the arc elasticities of demand for each mode of transport. Table 9 contains these arc elasticities for changes in rate, time in transit and reliability of 10 to 60 percent. In all cases, the elasticities are quite small for all attributes indicating that shipper demand for transport tends to be inelastic with respect to these transportation attributes. It appears grain elevators in the Ohio River Basin have few transportation alternatives and tend to be captive shippers. Consequently, most respondents either do not switch in response to a worsening in transportation conditions or decide to shut down.

Furthermore, the elasticities are remarkedly stable as attribute changes get larger. From a 10 percent change to a 60 percent change in transport attributes the arc elasticities remain generally constant.

Another interesting observation is that the estimated arc elasticities on rate and time-in-transit are relatively small when compared to previous estimates in similar studies. While these previous studies are based on similar surveys and comparable methodologies, estimates of shipper transportation demands are based on different geographies. This spatial variation provides one potential explanation for the differences in elasticity estimates. Shippers in the Ohio River Basin may have fewer shipment alternatives than similar shippers near the Upper Mississippi or Columbia-Snake rivers.

Moreover, changes in reliability are estimated to cause considerably larger changes in shippers' mode choices when compared to changes in rate and transit time. The underlying cause for these differences in the estimated arc elasticity is the notable difference in magnitude of the parameter estimates presented in Table 8. It appears that reliability is a key attribute across all three modes of transportation and that shippers are less responsive to changes in rate and time in transit than reliability. This is particularly true for shippers choosing truck.

Finally, a comparison across transport modes points to notable differences in elasticity estimates. With respect to changes in transport rate and time in transit, shippers utilizing truck are least responsive. Rail shippers are most likely to switch in response to rate increases, whereas barge shippers are more responsive to changes in time-in-transit. These barge shipments are also the longest in transit. With respect to worsening reliability all shippers demonstrate greater elasticities (still inelastic) irrespective of transport mode. In fact, shippers choosing truck transportation are most likely to switch in response to worsening reliability.

Table 9: Arc Elasticities

Rate	Percent Increase	Barge	Rail	Truck
	10	0.049	0.182	0.033
	20	0.049	0.183	0.033
	30	0.049	0.184	0.033
	40	0.05	0.185	0.033
	50	0.05	0.185	0.032
	60	0.05	0.186	0.032
Time	10	0.183	0.122	0.004
	20	0.189	0.124	0.004
	30	0.195	0.125	0.004
	40	0.201	0.126	0.004
	50	0.206	0.127	0.004
	60	0.211	0.127	0.004
Reliability	10	0.253	0.287	0.352
	20	0.26	0.292	0.342
	30	0.267	0.296	0.331
	40	0.274	0.3	0.318
	50	0.28	0.302	0.305
	60	0.285	0.303	0.29

5. Potential Limitations

As is the case with many studies that rely on survey data, there are a few noteworthy considerations. One of these commonly raised issues is the potential for behavioral biases in stated-preference survey responses. Applied in the context of this study, one may argue that respondents are more likely to state to remain with their initially chosen alternative in any of the SP experiments in order to reduce the cognitive costs of future questions. While this is a reasonable critique for surveys that contain a number of follow-up questions upon a respondent choosing to switch, this issue is less of a concern in this setting. In fact, shippers were asked to specify alternatives before being presented with the SP experiments and the only follow-up question arises when a shipper decides keep the status quo. Therefore, it is unlikely the respondents were affected by the consideration of the cognitive cost of future questions.

Another concern may be that expectations are not fully specified, and interpretations are subjective. For example, respondents may be uncertain about the duration of the shut-down option or have idiosyncratic interpretations of what it means to have changes in reliability. In the survey, changes in shipment attributes (i.e. rate, time, reliability) for each of the SP experiments is clearly described as a permanent change and option to the shut-down is labeled as “Go out-of-business”. Furthermore, the survey specifies that reliability is quantified through on-time delivery of shipments and that changes in reliability are to be interpreted as changes in the probability of on-time deliveries. Given the clarity of the survey, it is possible, yet difficult to imagine, that respondents have some uncertainty about the questions and options presented or idiosyncratic interpretations of reliability.

Relatedly, one potential concern may arise from the heterogeneity among shippers (i.e. grain elevators) in our sample and the influence of outliers. As discussed, the survey respondents include new (age<10 years) and old grain elevators (age>100 years), small (capacity<10,000 tons) and large shippers (capacity>100,000 tons), single- and multi-elevator firms, exporters and non-exporters. These differences may matter. For example, it is conceivable that larger firms have more bargaining power across all shipping characteristics and are therefore less sensitive to changes in any one of the shipping attributes. Reassuringly, restricting the estimation sample to grain elevators within the 10^{th} to 90^{th} percentiles of the observable characteristics yields quantitatively and qualitatively similar results.

6. Summary and Conclusions

Transportation is a critical component of agricultural markets. The flow of goods depends critically on the prices received for the product shipped, the rates, and other attributes associated with each mode and destination alternative. This reliance is an essential part of planning infrastructure, where the benefits of investment rest heavily on the responsiveness of demand to changes in rates.

In this study, we document survey evidence on originating shippers (grain elevators) in the Ohio River Basin. There are a large number of destinations where products can be sold, and there are major differences in shippers according to the modes available to make shipments. On average, shippers have been in business for a long period of time. Most are small shippers, but there are a number that are quite large in terms of tonnages shipped in a year and the storage capacity of the grain elevator. And, most elevators are single elevator firms, but, again, there are many elevators that are part of a firm that operates a multiplicity of grain elevators. Shipments from these elevators are dominated by truck movements to processing plants (ethanol, crushing, mills), but there are also a number that ship to river terminals or other elevators including export terminals.

We use these data to estimate transportation demands. The choice set for each elevator is solicited from the respondent along with the choice made. For each option, the survey solicits information prices received for the product transported, the transportation rate(s), the time in transit, and a measure of how reliable the mode is. These form the “revealed data”. Given shortcomings of revealed data e.g., small variation in the explanatory variables, we augmented the data with stated preference data wherein the stated preference data are the result of experiments on the revealed choice which helps to overcome a common criticism of stated preference data – that the stated preference experiments lack realism. We use a recently developed estimator that employs both types of data. This estimator corrects for any bias introduced by the fact that the stated preference information is constructed from the revealed choice made by the shipper. That is, it accounts for the transfer of the unobserved attributes of the revealed choice onto the SP experiments. We find that each of the attributes has a statistically significant effect on the choices made by shippers. And we find that elasticities generated are relatively small in magnitude, pointing to inelastic demands. That is, the demand functions appear to be reasonably steep and point to a large degree of captive shippers (i.e., shippers that do not switch to alternatives even for large changes in the attributes). While this result points to relatively large benefits to infrastructure investments, there are limits. The analysis also includes the incorporation of the option of no longer shipping (i.e., shutting down). This finding has been a consistent theme in related research. In the present case, the option to shut down is explicitly represented in the choice model. Hence, attributes, particularly rates, cannot increase without bound because eventually shippers will opt out of the market. Our findings suggests that the option to shut down is very important in gauging the responsiveness of demand to rates.

References

Babcock, M.W. and Gayle, P.G., 2014, March. Specifying and Estimating a Regional Agricultural Railroad Demand Model. *Journal of the Transportation Research Forum* (Vol. 53, No. 1).

Bergantino, A., Bolis, S., 2006. An adaptive conjoint analysis of freight service alternatives: Evaluating the maritime option. In Schintlerm, L., Reggiani, A. (Eds.), Methods and Models in Transport and Telecommunications: Cross-Atlantic Perspectives. Springer Verlag, Berlin.

Clark, C., H.T. Naughton, B. Proulx, and P. Thoma. 2005. "A survey of the freight transportation demand literature and a comparison of elasticity estimates." IWR Report 05-NETS-R-01. Available at https://planning.erdc.dren.mil/toolbox/library/IWRServer/05-NETS-R-01.pdf.

Caussade, S., de Dios Ortúzar, J., Rizzi, L.I. and Hensher, D.A., 2005. Assessing the influence of design dimensions on stated choice experiment estimates. Transportation research part B: Methodological, 39(7), pp.621-640.

Dybing, A. G. 2002. *Estimation of the Demand for Grain Transportation in North Dakota*. Unpublished Master's Thesis: North Dakota State University.

Fitzsimmons, E.L., 1981. A statistical sketch of the demand for rail transport of grain and soybeans. *Transportation Journal*, pp.59-65.

Fowkes, A., Shinghal, N., 2002. The Leeds adaptive stated preference methodology. In: Daniels, R. (Ed.), Freight Transportation Demand and Stated Preference Experiments. FrancoAngeli, Milan, Italy.

Fowkes, A.S., Firmin, P.E., Tweddle, G. and Whiteing, A.E., 2004. How highly does the freight transport industry value journey time reliability—and for what reasons?. *International Journal of Logistics: Research and Applications*, *7*(1), pp.33-43.

Friedlaender, A.F. and Spady, R.H., 1980. A derived demand function for freight transportation. *The Review of Economics and Statistics*, pp.432-441.

Guevara, C.A. and S. Hess. 2019. A control-function approach to correct for endogeneity in discrete choice models estimated on SP-off-RP data and contrasts with an earlier FIML approach by Train and Wilson. *Transportation Research Part B: Methodological*, 123:224–239.

Henrickson, K.E., 2011. Spatial competition and geographic grain transportation demand on the Mississippi and Illinois rivers. *Applied Economics*, *43*(10), pp.1257-1269.

Hensher, D., 2004. Accounting for stated choice design dimensionality in willingness to pay for travel time savings. Transportation Research Part B 38, 425–446.

Hensher, D., 2006. How do respondents process stated choice experiments? attribute consideration under varying information load. Journal of Applied Econometrics 21, 861–878.

Hensher, D., Greene, W., 2003. Mixed logit models: state of practice. Transportation 30, 133–176.

Inaba, F.S. and Wallace, N.E., 1989. Spatial price competition and the demand for freight transportation. *The Review of Economics and Statistics*, pp.614-625.

Khakdaman, M., Rezaei, J., & Tavasszy, L. A. (2020). Shippers' willingness to delegate modal control in freight transportation. *Transportation Research Part E: Logistics and Transportation Review*, *141*, 102027.

Kalahasthi, L., Holguín-Veras, J., & Yushimito, W. F. (2022). A freight origin-destination synthesis model with mode choice. *Transportation Research Part E: Logistics and Transportation Review*, *157*, 102595.

Keya, N., Anowar, S., & Eluru, N. (2019). Joint model of freight mode choice and shipment size: A copula-based random regret minimization framework. *Transportation Research Part E: Logistics and Transportation Review*, *125*, 97-115.

Miklius, W., Garrod, P.V. and Casavant, K.L., 1976. Estimation of demand for transportation of agricultural commodities. *American Journal of Agricultural Economics*, *58*(2), pp.217-223.

Miljkovic, D., Price, G.K., Hauser, R.J. and Algozin, K.A., 2000. The barge and rail freight market for export-bound grain movement from midwest to Mexican Gulf: an econometric analysis. *Transportation Research Part E: Logistics and Transportation Review*, *36*(2), pp.127-137.

Oum, T.H., 1979a. A warning on the use of linear logit models in transport mode choice studies. The Bell journal of economics, pp.374-388.

Oum, T. H. 1979b. A Cross-Sectional Study of Freight Transport Demand and Rail-Truck Competition in Canada. *The Bell Journal of Economics,* 10(2): 463-482.

Oum, T. H. 1979c. Derived Demand for Freight Transport and Inter-modal Competition in Canada. *Journal of Transport Economics and Policy,* 13(2): 149-168.

Oum, T.H. 1989. "Alternative Demand Models and Their Elasticity Estimates." *Journal of Transport Economics and Policy*, 5:163–87.

Oum, T.H., W.G. Waters II, and J.S. Yong. 1992. "Concepts of Price Elasticities of Transport Demand and Recent Empirical Estimates: An Interpretative Survey." *Journal of Transport Economics and Policy*, 26(2):139–69.

Rose, J., 2005. Designing efficient stated choice experiments. PhD. thesis, School of Business, University of Sydney.

Shinghal, N., Fowkes, A., 2002. Freight mode choice and adaptive stated preferences. Transportation Research, Part E: Logistic and Transportation Review 38, 367–378.

Sitchinava, N., W. Wilson, and M. Burton. 2005. "Stated Preference Modeling of the Demand for Ohio River Shipments." IWR draft report. Available at http://www.nets.iwr.usace.army.mil/docs/other/OhioStatedPreferenceDemand.pdf

Train, K., and W.W. Wilson. 2004. "Shippers' Responses to Changes in Transportation Rates and Time: The Mid-American Grain Study." IWR Report 04-NETS-R-02. Available at https://planning.erdc.dren.mil/toolbox/library/IWRServer/04-NETS-R-021.pdf.

Train, K. and Wilson, W.W. 2008. Estimation on stated-preference experiments constructed from revealed preference choices. *Transportation Research Part B: Methodological*, 42(3):191–203.

Train, K. and Wilson, W.W. 2008b. "Transportation Demand and Volume Sensitivity: A Study of Grain Shippers in the Upper Mississippi River Valley." *Transportation Research Record*, No. 2062:66–73.

Train, K. and Wilson, W.W. 2019. "Demand for the Transportation of Agricultural Products: An Application to Shippers in the Upper Mississippi and Illinois River Basins". Report to the USACE.

Wilson, W.W., 1984. Estimation of modal demand elasticities in grain transportation. Western Journal of Agricultural Economics, pp.244-258.

Wilson, W.W., Wilson, W.W. and Koo, W.W., 1988. Modal competition in grain transport. Journal of Transport Economics and Policy, pp.319-337.

Winston, C. 1983. "The Demand for Freight Transportation: Models and Application." *Transportation Research – Part A,* 17A:419–427.

Yu, T.H. and Fuller, S.W., 2005. The measurement of grain barge demand on inland waterways: a study of the Mississippi River. In Journal of the Transportation Research Forum Vol. 44, No. 1424-2016-117976, pp. 27-39).

Yu, T., Zhang, J., and Fuller, S. W. (2006). Freight Transportation Demands on the Upper Mississippi and Illinois Rivers. *Journal of the Transportation Research Forum,* 45(2): 119-136.