A-3. METHODOLOGICAL APPENDIX

A number of the variables used to estimate railroad pricing behavior are obtained directly from the Carload Waybill Sample and appear in the specified model without manipulation. However, a number of the relevant variables are constructed from the waybill data and/or other data sources. A precise and detailed discussion of this latter group of variable is provided below.

Distance-To-Water Measure

Obviously, the most important variable within the context of this analysis is the shipment distance to water measure(s) included in the estimated models. From a purely theoretical vantage, both distance of a shipment's origin to the nearest navigation resource and distance to water at the destination should impact the desirability of the barge alternative. In practice, however, the relative importance of the distance to water at the origin and the distance to water at the destination is an empirical matter. In some cases, most or all origins may be at or near a navigation resource, so that it is the destination distance to water which is the most important determinant of railroad pricing. It is equally possible to encounter situations in which the terminal distance to water is unimportant relative to the origin distance to the nearest waterway.

As the text indicates, the relationship between distance to water and observed rates is discontinuous over the full range of shipment distances. Specifically, at some critical distance from the water, available navigation ceases to have any effect on rail rates. For estimation purposes, this critical distance is reflected by two dummy variables, $OCDUM_i$ and $TCDUM_i$. The value of the former variable is equal to one if the origin distance to water is less than the critical distance beyond which water has no impact and zero otherwise. Similarly, $TCDUM_i$ takes on a value of one if the destination distance to water is less than the appropriate critical distance and zero otherwise. In order to account for a full range of possibilities, the estimation process for each commodity began with the same specification which is summarized by Equation A1 below:

$$\text{(A1)} \quad RTM_i = \delta_1 + \delta_2(OD2W_i) \times (OCDUM_i) + \delta_3(TD2W_i) \times (TCDUM_i) + \delta_4(OCDUM_i) + \delta_5(TCDUM_i) + b\,\mathbf{X} + \varepsilon_I$$

where RTM_i is the revenue per ton-mile, $OD2W_i$ is the origin distance to water, $TD2W_i$ is the destination distance to water, b is a vector of regression coefficients, and **X** is a vector of other independent variables. This specification allows for either or both of the relevant distances to water to affect the observed railroad rate. If either combination of dummy variable and interaction term is jointly insignificant at the ten percent level, that combination was dropped from the model specification and the model was re-estimated. If available water transportation has the assumed dampening impact on railroad rates, the signs for the two interaction terms are positive and the signs of the two dummy variables are negative.

In order to determine the appropriate critical distance, the model described by Equation (A1) was estimated iteratively. At each iteration, the value defining each dummy variable was incremented by five miles. When the joint probability that an interaction term and its associated dummy variable are both different from zero was maximized, that particular distance was fixed while the routine continued to increment the definition of the remaining dummy variable until the joint probability for that interaction/dummy pair was also maximized. At that point, the first pair to converge was re-estimated to verify its stability and the process was continued until a stable pair of probability maximizing distances was obtained.

The actual distances are calculated as straight-line distances from the most active business location in the county of origin/termination to a major general commodities port.[1] Finally, because trans-shipment imposes fixed costs which must be averaged over the entire shipment distance, all distance to water measures were weighted by the total shipment distance.

Railroad Market Concentration

[1] The most active business location within each county is defined as that city or town with the greatest number of business addresses.

In past investigations, we have used a number of different measures to capture the importance of intramodal railroad competition as a determinant of observed rates.[2] In this investigation, the richness of the waybill data allowed us to construct a new measure which seems to improve our ability to account for this competition. In the analysis $RRCON_{ij}$ is defined as the product of the originating carrier's market share at origin ***i*** with the delivering carrier's market share at destination ***j***. This specification treats the multi-line production of railroad transportation as a vertical relationship and, as with any such vertical relationship, market power at any stage in the process is sufficient to generate higher prices.

Route Density.

In the absence of truly reliable route information, it is nearly impossible to fully account for the effects of traffic density on railroad costs (and rates). For the purposes of this analysis, a density is calculated for each carrier or combination of carriers serving a particular state-to-state origin-destination pair. The value of this calculation is equal to the sum of transported tons across all commodities divided by the mean distance for the carrier(s)' movements over the particular origin and destination pair. The data support the construction of an analogous measure over smaller geographic units (either BEA areas or counties), but the route structures of most carriers seem to indicate that the state-to-state measure is preferable.

Car Ownership

Unlike past efforts, these estimations explicitly account for whether the equipment used in a particular movement is owned by a railroad or by the customer (or some third party). Table A2 contains the list of railroads reporting marks used to determine whether or not a particular car is a system car.

Table A2

[2] Previous measures included the number of carriers offering service between an a particular origin-destination pair and a Herfindahl-Hirschmann type statistic calculated over a particular market.

ALS	AM	AKMD	ALM	ATSF	SFRC	BAR	BM	BN	BNFE
CBQ	CS	FWD	GN	NP	BBN	RBBQ	RBCS	RBW	SLSF
SFE	WHI	CN	BCNE	CAN	CNIS	CVC	DWC	NAR	CV
CGW	CMO	FDDM	LM	MSTL	CC	CAGY	CR	BA	BCK
CNJ	CLW	EL	ERIE	MGA	NH	NYC	PAE	PC	PRR
RDG	RR	TDC	CP	CPAA	CPI	CPT	DA	NJ	THB
CSXT	ACL	AWP	BO	CO	CRR	GA	LN	MON	NC
RFP	SAL	SBD	SCL	WA	WM	DME	DH	DHNY	DRGW
EJE	ELS	FEC	GVSR	GTW	DTI	DTS	IC	CIW	GMO
ICG	IHB	IAS	KCS	CTIE	GNA	MSRC	KYLE	MSDR	MP
ARDP	ARMH	ARNW	BKTY	CHTT	DKS	MI	MKT	MKTT	OKKT
TP	MRL	NS	NW	PWV	SA	SOU	TAG	VGN	NOKL
PAL	PPU	SRN	SLR	SSW	SOO	MILW	MNS	SP	SPFE
GMSR	SR	UP	SI	TNM	SPFE	WP	WPMW	WE	WC

Carrier Dummy Variables

In addition to the other right-hand-side variables, each estimation contained a set of zero/one dummy variables designed to indicate a specific carriers participation in the shipment. Each of these variables assumes a value of one if the particular carrier originated or terminated the shipment and zero otherwise.[3]

[3] This method fails to represent the participation of a bridge carrier which neither originates nor terminates the shipment. However, given that the mean number of carriers is significantly less than two for each of the commodities and that bridge carriers have a diminished influence over price, we do not feel this is inappropriate.