

August 31, 2018

William Brennan, Director
Office of Economics, Environmental Analysis, and Administration
Surface Transportation Board
395 E Street SW
Washington, DC 20423
Subject: Waybill Sample Data Request

ENTERED
Office of Proceedings
September 4, 2018
Part of
Public Record

Dear Mr. Brennan:

Please consider this a request of access to the Unmasked Carload Waybill Statistics from 1983 to 2017. This request will allow continuation of academic research on the railroad industry by Professor Wesley W. Wilson of the University of Oregon and Frank A. Wolak of Stanford University (Principle Investigators) which was initiated under previous access agreement. This is a modest update from our previous letter dated August 25, 2016. It allows access to more recent data for updating and extending our research.

The Unmasked Carload Waybill Statistics are the best data available, and the collection of these data from other sources is not only burdensome, it is likely impossible. Below we provide the purpose for which access is needed, a description of specific waybill data or fields that are actually required (including pertinent geographic areas), and a detailed justification as to why the specific waybill data are needed. I note that we are requesting the entirety of the unmasked waybill. Most of the fields are required for the research and the few that are not are superfluous do not mandate special programming either by the STB or by the principle investigators who have done extensive work with the same data previously. We also fully understand and will comply with the confidentiality standards (the three shipper rule, the three-FSAC rule) and subject any research output using these data to the Director of the Office of Economics for confidentiality compliance review.

Access to the data is needed to continue a line of research is to examine the determination of rates, costs, and service across railroads, time, and geography. As is well recognized in the economics literature, a massive number and wide variety of outputs are produced by railroads as delineated by commodity movements and origin-destination pairs and types of railroad configurations employed i.e., single car, multicar, and unit trains. Analysis of costs is complicated by the sheer dimensionality of the set of outputs, the fact that most costs are common or joint costs, the presence of both economies of density and possibly economies of scale as well as cost complementarities across different outputs. Analysis of rates is complicated by each of these factors as well as the fact that there are at least three different pricing regimes (tariff rates, contract rates, and exempt rates) and differential levels of competitive alternatives available to shippers. Generally, the set of rates must compensate not only for the costs of the shipments, but also the costs of providing and maintaining the network to ensure that railroads are revenue adequate

Over the last several years, the authors have developed a set of programs that integrate the R-1 data to the Waybill data and have augmented the data with data reflecting the options available to shippers and the spatial locations of shippers along with intramodal and intermodal options available to shippers as well as link-specific information on the railroad networks. These data are the most complete, to our knowledge, available to conduct meaningful research on the railroad industry. To this end, we have identified projects which are outlined below with a list of fields in the waybill necessary to complete the

work. These are: 1) An examination of the determinants rail rates; 2) The determinants of railroad and shipment costs; 3) Factors leading to rail network adjustments (retirements, expansions, and quality improvements through investments; and 4) Investigation spatial relationships in international trade and in coal markets.

PLANNED RESEARCH

Rates and Revenue Adequacy: There are a number of different studies of the determinants of railroad rates. Some of these use aggregate rate levels, while other use some version of rates taken from the waybill data. Generally, controls such as distance, shipment size, the number of interchanges etc. are introduced as cost variables, while controls such as distance to waterways and measures of railroad concentration are introduced to control for differing competitive factors. We seek to estimate more flexible models by introducing richer measures of costs and competitive factors for a shipment, while accounting for the pricing regime (tariff, contract, and exempt). We then plan to use these models to examine alternative strategies for identifying rates that are significantly different from rates charged for observably similar shipments and for alternative strategies for railroads, shippers and regulators to identify solutions for these misaligned rates. The analysis cannot be done without access to the unmasked waybill and is key to understanding the determinants of shipment rates and the ability of the rate regimes to allow railroads to be revenue adequate.

Railroad Costs: Estimation of railroad costs has a long history. Different analyses and models have been conducted to uncover both scale and density measures, productivity measures, as well as to estimate movement specific costs. Each of these tasks face serious methodological challenges because of the number and variety of products a railroad provides. Most of the difficulties are directly the result of the fact that the number of outputs produced by a railroad is far too large to be feasible to estimate a multi-product cost function given the data available without extreme assumptions. This is a long-recognized problem, and most research has used aggregate measures e.g., total ton-miles or some aggregation within specific, but limited, product classifications e.g., agricultural goods, or production type e.g., bulk, classifications. The problem of dimensionality necessitates such aggregation. We seek to reduce the dimensionality issue by following recent advances in modeling demands products with an enormous number of substitutes. In such cases, economists have mapped the product space into an attribute space to reduce the number of dimensions that plague econometric analysis. In the case of railroads, there are a relatively small set of attributes available in the waybill that allow estimation of costs through attributes rather than separate outputs e.g., origin estimation-commodity. The attributes are readily available in the waybill and the cost information available in the R-1 reports. This analysis cannot be completed without access to the waybill.

Network Adjustments: The third theme of the research program relates to railroad networks. Since deregulation of railroads there has been a massive reduction in miles of road and locations served by Class 1 carriers as firms abandoned or sold line segments. Using companion data, we are able to identify stations and line segments that are no longer served by a specific railroad and the disposition of the station or line to an alternative carrier e.g., a short-line or in terms of cessation of any railroad service. With the waybill, we can estimate that total movements associated not only with the station but also along the line segments. This allows exit/sale decisions to be mapped into measures of station size, economies of density along the line segment, as well as line characteristics, revenues, firm characteristics and demand characteristics. The result will allow a novel examination of the determinants of rail networks and how they evolve over time in response to market conditions, firm characteristics, and policies that hit the rail industry directly as well as indirectly.

Spatial Relationships: The fourth theme of this research program relates to the interdependence of rail to other markets. In the research, we plan to integrate models of international trade to rail, barge and truck markets, and assess the effects of rail on trade flows. In a second planned research area, we plan to examine the change in environmental regulation and natural gas prices on the coal market and on the transportation of coal across spatial boundaries. In a third planned research area, we examine the relationship between oil prices, transportation rates, and transportation flows of crude oil.

WAYBILL DATA AND VARIABLES

We request data from 1983-present.

RESEARCHERS

We request permission for four researchers. These include: Wesley W. Wilson and Frank A. Wolak (principal investigators), Kenneth D. Boyer, Brett Garcia, Tobias Sytsma, and Curtis Dlouhy.

The principal investigators are Wesley Wilson and Frank Wolak. They are Professors in their respective departments, and we have attached the vitas of each to this request. Kenneth Boyer is a Full Professor in his department, while Brett Garcia, Tobias Sytsma, and Curtis Dlouhy are graduate students. The latter four are potential collaborators, and as such, will need access to intermediate files for research facilitation.

The principal investigators have worked with the waybill for years and will adhere to the confidentiality standards of the waybill. The waybill will be used solely for academic research purposes.

If you have any questions, comments or concerns, please contact us. The contact information is given below.



Sincerely,

Wesley W. Wilson
Department of Economics
University of Oregon
Eugene, OR 97403-1285
wwilson@uoregon.edu
tele: (541) 346-4690
fax: (541) 346-1243



Frank A. Wolak
Department of Economics
Stanford University
Stanford, CA 94305-6072
wolak@zia.stanford.edu
tele: (650) 724-1712
fax: (650) 724-1717