

## Creation of the Public Use Waybill Sample

The Public Use Waybill Sample (PUWS) is a non-proprietary version of the STB Carload Waybill Sample. The STB collects the data under the requirements that all US railroads that terminate more than 4,500 revenue carloads must submit a yearly sample of terminated waybills.

The waybills are sampled under two different plans. The “MRI” plan for data delivered on cartridges stratifies sampled waybills into five different levels of sampling frequency depending on the number of carloads on the waybill. Waybills representing larger number of carloads are sampled more frequently. The second technique called the “Ex Parte 385” plan allows manual sampling of waybills, typically for smaller railroads, and is stratified into three different levels of sampling frequency. The entire sample of waybills is then expanded using appropriate multipliers for each sampling level to represent total US rail movements in that year.

Use of the waybill data is subject to some qualifications. As with any sample, some portions of the total population are better represented than others.

Since the full Carload Waybill Sample contains specific waybill information such as origin and termination freight station, junction points, and rail carrier identification, it is not suitable for public release. As an alternative, the Public Use Waybill Sample has been created from the original full sample by elimination station and carrier information. Origin and termination points are reported by Business Economic Area (BEA) and junction points are reported by state or province, rather than by freight station or city name.

Additionally, some waybill records are excluded from the PUWS. The PUWS only contains rail freight movements for commodities that were handled by at least three freight stations in the US. If a 5-digit commodity was not handled by at least three Freight Station Accounting Codes (FSACs) nationwide, the record is rejected from the PUWS.

Commodities (with the exception of munitions data) are identified at the 5-digit STCC level. Because of the sensitive nature of the munitions data, this information is reported at the 2-digit STCC level (STCC 19) and no geographic coding for these records is included.

The use of Business Economic Areas (BEAs) in the PUWS is subject to the “three-FSAC rule”. This rule was adopted to protect against any disclosure of competitively sensitive waybill data in the Public Use file. Under this approach, a BEA is only reported if there is activity for at least three FSACs on one railroad for a given commodity within that BEA, or if there are at least two more FSACs with activity than there are railroads in that BEA for a given commodity. Records that do not pass the three FSAC rule are still included, but without any geographic coding. Intermediate junction data is shown only when both

the originating and the terminating BEAs pass these criteria. Only about 45 to 50% of the total waybill records have full geographic data.