APPENDIX C

SELECTION OF CALIFORNIA CASE RECORDS FOR MATCHING

Summaries of the steps involved in the case selection process (along with the cases retained / excluded at each step) for preparing the California WC and MSHA data files are presented in the following tables.

California WC File Case Selection

Step	Dropped	Case Count
Step 1 – Verify injury dates from 2005 – 2009, inclusive	17	9080
Merge the NAICS tab with the SIC tab and deduplicate retaining NAICS case if duplicate JCN exists	651	8429
Number of cases dropped in Step 1	**668**	
Number of cases available for Step 2		**8429**
Step 2 – Merge with Tab 3 (Class Code) tab and deduplicate again		
Retaining the case from step 1 if duplicate exists		
Number of duplicate cases dropped in Step 2	**214**	
Number of cases available for Step 3		**8215**
Step 3 – Merge with Tab 4 (Mining in string var) tab and deduplicate again		
Retaining the case from step 2 if duplicate exists		
Number of cases dropped in Step 3	**2132**	
Number of cases available for Step 4		**6083**
Step 4 – Select cases that occurred in California		
Select out cases not from California CA zip codes are 900xx – 961xx. Exclude cases that do not have a valid CA zip code, including those with missing or invalid zip code.		
Number of cases dropped in Step 4	**224**	
Number of cases available for Step 5		**5727**
Step 5 – Select cases which are valid injuries and eliminate illnesses and non-physical harms		
Number of cases dropped in Step 5	**660**	
Number of cases available for Step 6		**5067**
Step 6 – Eliminate cases with conflicting multiple codes (NAICS, SIC, CLASS CODE, Industry Description) which appear not to be mining		
Number of cases dropped in Step 6	**1575**	
Number of cases available for analysis		**3492**

California MSHA File Case Selection

Step	Dropped	Case Count
Cases in the MSHA employee (aii) file 2005 – 2009	--	**68,113**
Step 1 – Select cases of accident and injury (where variable 'aii' is 21 or less)		
Cases with aii values over 21	1,625	
Number of cases dropped in Step 1	**1,625**	
Number of cases available for Step 2		**66,488**
Step 2 – Select cases that occurred in California (using variable 'state')		
Cases in states other than California	64,462	
Number of cases dropped in Step 2	**64,462**	
Number of cases available for analysis		**2,026**