Documentation for "Demand Conditions and Worker Safety: Evidence from Price Shocks in Mining"

Kerwin Kofi Charles   Matthew S. Johnson   Melvin Stephens, Jr.   Do Q Lee

February 3, 2021

1 Data Documentation

This document describes the replication archive for the above paper. This archive contains the public-use data files used in the paper and the programs to run the analysis. Two main datasets:

1. *MSHA Mining Data.* We assemble a mine-level panel dataset containing quarterly employment, injuries, MSHA inspections, violations of MSHA regulations, corporate ownership, and other mine attributes from the Mine Safety and Health Administration's (MSHA's) open government data sets. The datasets were downloaded from the MSHA website on September 2015, and can be accessed at: https://enforcedata.dol.gov/views/data_summary.php

2. *Price Data.* We assemble a commodity-level panel from four sources.

   Our main source is the London Metal Exchange (LME), the world's largest market for financial contracts on metals. For commodities not covered in the LME database, we use prices from the International Monetary Fund (IMF) Primary Commodity Prices database. To further expand our set of commodities, we use Datastream to access the quarterly cash price 1983-2015 for metals in the MSHA database. The prices from Datastream are derived from a number of sources including Standard & Poors and Thompson Reuters. Finally, for commodities not included in any of the databases above, we obtain additional data from Bloomberg.

Charles: Yale School of Management, kerwin.charles@yale.edu; Johnson: Duke University, matthew.johnson@duke.edu; Stephens: Univerity of Michigan, mstep@umich.edu; Lee: New York University, dql204@nyu.edu.

The price data sourced from Datastream and Bloomberg are proprietary. Researchers interested in reproducing the analysis will need to obtain a subscription for these data sets.

We also draw from three other datasets:

1. *World Mining Data.* For Table 3. Contains data about the Herfindahl-Hirschman Index (HHI) for the U.S. and the world. Data comes from the International Organizing Committee for the World Mining Congresses, 2014.

2. *Production Data (USGS).* For Table 3. Contains data on production for each commodity (in various units). Publicly available and longer time series coverage, but only for gold and silver.

3. *Production Data (Datastream).* For Figure B.1. Commodity-level panel of production data (in various units). Proprietary.

2   Code Documentation

The code is stored in `code/`

1. Contains all the programs needed to go from the raw data files to the results in the paper.

2. `code/main.do` is the master do-file. Running this program executes the replication by calling the remaining do-files in this folder. The do-files are written for Stata version 14 or above.

3. `code/ado/` contains any additional Stata packages called by `code/main.do`.

4. `code/log/` contains log files in `.txt` format.

Input to the programs in `data/`

1. `data/msha/` contains the raw data files downloaded from MSHA.

2. `data/price` contains the price data downloaded from various sources.

3. `data/production/` contains data on production.

4. `data/bmwfw/` contains data on the Herfindahl-Hirschman Index (HHI) for each commodity.

Data Preparation:

1. `code/msha_qtrly_oprtr_emplymnt.do` prepares the MSHA operator employment data

2. `code/msha_qtrly_cntrctr_emplymnt.do` prepares the MSHA contractor employment data

3. `code/msha_accident.do` prepares the MSHA accidents data (injuries and illnesses)

4. `code/msha_violation.do` prepares the MSHA violations data

5. `code/msha_inspection.do` prepares the MSHA inspections data

6. `code/msha_operator_history.do` prepares the MSHA operator history data

7. `code/msha_controller_history.do` prepares the MSHA controller history data

8. `code/msha_contractor_history.do` prepares the MSHA contractor history data

9. `code/price.do` prepares the commodity price data

10. `code/datastream_production.do` prepares the production data from datastream

11. `code/mine_panel_for_analysis.do` prepares the main dataset for the analysis.

12. `code/mine_panel_for_siblings_analysis.do` prepares a version of the main dataset with information on corporate siblings

Tables: `code/table#.do` with # corresponding to the table number in the draft.

Figures: `code/figure#.do` with # corresponding to the figure number in the draft.

Output to the programs:

1. `figures/` contains the figures in the paper in `.eps` format.

2. `tables/` contains the tables in the paper in `.tex` format.