Central Appalachian Mine Closures

https://github.com/hgs13/EDA_Final_Project_2020

Hannah Smith

Contents

1 Rationale and Research Questions **5**

2 Dataset Information **6**
2.1 Data Origin Location 6
2.2 Data Wrangling Methods 6
2.3 Data Exploration 7

3 Exploratory Analysis **11**
3.1 Visual Data Exploration of Harlan and Dickenson Data 11

4 Analysis **20**
4.1 What is the relationship between coal production and number of employees in coal mines in Harlan and Dickenson Counties? 20
4.2 How does coal production differ between all mines in Harlan County compared to Dickenson County? 20
4.3 How does the number of employees in the mines of Harlan County compare to Dickenson County? 20
4.4 Is the total number of employees for all mines in a boom-bust and a bust-bust county in the years 2000 to 2011 a significant predictor of annual coal production for all mines? 20

5 Summary and Conclusions **24**

6 References **25**

List of Tables

List of Figures

1 Total Annual Labor Hours for Mining in Harlan and Dickenson Counties by Mine . . . 12
2 Total Annual Labor Hours for Mining in Harlan and Dickenson Counties by County . . . 13
3 Total Annual Coal Employment in Harlan and Dickenson Counties by Mine . 14
4 Total Annual Coal Employment in Harlan and Dickenson Counties by County 15
5 Total Annual Coal Production in Harlan and Dickenson Counties by Mine . 16
6 Total Annual Coal Production in Harlan and Dickenson Counties by County 17
7 Total Annual Coal Production versus Total Employment in Harlan and Dickenson Counties by Mine . . . 18
8 Total Annual Coal Production versus Total Employment in Harlan and Dickenson Counties by County . . . 19
9 Total Annual Coal Production in Harlan and Dickenson Counties . . . 20
10 Total Number of Employees in Harlan and Dickenson County Mines, 2000 to 2011 . . . 21
11 Annual Coal Production Compared to Total Employees in Harlan and Dickenson Counties, 2000 to 2011 . . . 22

1 Rationale and Research Questions

This project seeks to analyze preliminary data that may affect the migration patterns of residents of coal counties in central Appalachia from the years 2000 to 2011. There is little to no research exploring migration patterns in Appalachia during this time period, especially as it relates to migration. Furthermore, the energy sector is transitioning to greener energy, which further impacts coal production, especially in Appalachia.

Two "coal counties" serve as a case study for migration in central Appalachia: Harlan County, Kentucky and Dickenson County, Virginia. Harlan County serves as a "Boom-Bust" county, meaning the county experienced a coal production boom in 2000 and a coal production bust in 2010. Dickenson, on the other hand, is a "Bust-Bust" county, where the county experienced a coal production bust in 2000 and did not recover by 2010.

An important portion of understanding migration in Appalachia is investigating the mine operations within the coal counties. As such, this project creates and analyzes a visualization of mine production and employment in Harlan and Dickenson counties between the years 2000 and 2011. The data analysis of the coal production and the number of employees in coal mines in each county could be indicative of migration patterns within the central Appalachian region.

The main questions for this research study are as follows:

1. What is the relationship between coal production and number of employees in coal mines in Harlan and Dickenson Counties?

a. How does coal production differ between all mines in Harlan County compared to Dickenson County?

b. How does the number of employees in the mines of Harlan County compare to Dickenson County?

c. Is the total number of employees for all mines in a boom-bust and a bust-bust county in the years 2000 to 2011 a significant predictor of annual coal production for all mines?

2 Dataset Information

2.1 Data Origin Location

The data in this repository is coal mine data compiled by the Coal and America Bass Connections team at Duke University from the Energy Information Administration (EIA) online database throughout the fall of 2019. The data in this repository is coal mine data compiled by the Coal and America Bass Connections team at Duke University from the Energy Information Administration (EIA) online database throughout the fall of 2019.

"This report is mandatory under the Federal Energy Administration Act of 1974 (Public Law 93-275). Failure to comply may result in criminal fines, civil penalties, and other sanctions as provided by law. Title 18 USC 1001 makes it a criminal offense for any person knowingly and willingly to make to any Agency or Department of the United States any false, fictitious, or fraudulent statements as to any matter within its jurisdiction.

All coal mining companies that owned a mining operation which produced 25,000 or more short tons of coal during the reporting year must submit form EIA-7A, except for anthracite mines. All anthracite mines that produced 10,000 or more short tons during the reporting year must submit form EIA-7A. Standalone facilities (e.g., preparation plant/tipple/loading dock/train loadout) that worked 5,000 or more hours must submit the EIA-7A. Submit a separate form EIA-7A for each mining operation and standalone facility that meets the reporting criteria.

The U.S. Energy Information Administration's (EIA) Form EIA-7A, Annual Survey of Coal Production and Preparation, collects coal production data from U.S. coal mining companies. This includes information on the type and status of coal operations, characteristics of coalbeds mined, recoverable reserves, productive capacity and the disposition of coal mined which provides Congress with basic statistics concerning coal supply. These data appear in the Annual Coal Report, the Quarterly Coal Report, the Monthly Energy Review, and the Annual Energy Review. In addition, the EIA uses the data for coal supply analyses and in short-term modeling efforts, which produce forecasts of coal supply and prices requested by Congress. The forecast data also appear in the Short-Term Energy Outlook and the Annual Energy Outlook." (EIA, Form EIA-7A)

Therefore, the data used in this project should be timely and accurate. Furthermore, coal production during the decade this project explores was extremely variable. As such, there should be no exclusion of outliers in this report.

2.2 Data Wrangling Methods

For this project, the data was imported as two separate .csv files for each county of interest (Harlan, Kentucky and Dickenson, Virginia). From there, each column was coerced into the appropriate vector format. The two data frames were then joined into one data frame to create ease of access. Then, irrelevant columns were removed so that only the year, mine name, mine state, mine county, mine status, operation type, operating company, average employees, and labor hours were included. After this, only active mines were selected for, as

inactive mines would skew the data. After grouping the remaining data by year and county, the total production, total employed, mean production, and mean empoyment were found for each county by year.

2.3 Data Exploration

Data exploration of the Harlan and Dickenson County raw data files

```
dim(harlan.raw)
```

```
## [1] 509  15
```

```
dim(dickenson.raw)
```

```
## [1] 194  15
```

```
str(harlan.raw)
```

```
## 'data.frame':    509 obs. of  15 variables:
##  $ year                     : int  2011 2011 2011 2011 2011 2011 2011 2011 2011 2011
##  $ mine.name                : Factor w/ 126 levels "# 1","# 2","# 3",..: 34 104 122 4
##  $ mine.state               : Factor w/ 1 level "Kentucky (East)": 1 1 1 1 1 1 1 1 1
##  $ countystr                : Factor w/ 1 level "Harlan": 1 1 1 1 1 1 1 1 1 1 ...
##  $ mine.basin               : Factor w/ 1 level "Appalachia Central": 1 1 1 1 1 1 1 1
##  $ mine.status              : Factor w/ 6 levels "Active","Active, men not working, n
##  $ mine.type                : Factor w/ 1 level "Underground": 1 1 1 1 1 1 1 1 1 1 ..
##  $ company.type             : Factor w/ 3 levels "Contractor","Indepedent Producer Op
##  $ operation.type           : Factor w/ 2 levels "Mine only","Preparation Plant": 1 2
##  $ operating.company        : Factor w/ 102 levels "A & M Coal Co Inc",..: 54 64 37 3
##  $ operating.company.address: Factor w/ 92 levels "1160 Jackson Dr, Paris, TN 38242",
##  $ production.stons         : Factor w/ 400 levels "0","1,013,140",..: 248 1 1 1 1 1
##  $ average.employees        : int  110 25 22 20 2 9 127 2 35 4 ...
##  $ labor.hours              : Factor w/ 468 levels "","1,211","1,692",..: 174 243 365
##  $ ARC                      : int  1 1 1 1 1 1 1 1 1 1 ...
```

```
str(dickenson.raw)
```

```
## 'data.frame':    194 obs. of  15 variables:
##  $ year                     : int  2011 2011 2011 2011 2011 2011 2011 2011 2011 2011
##  $ mine.name                : Factor w/ 56 levels "#1","#2","#44",..: 37 16 36 32 11
##  $ mine.state               : Factor w/ 1 level "Virginia": 1 1 1 1 1 1 1 1 1 1 ...
##  $ countystr                : Factor w/ 1 level "Dickenson": 1 1 1 1 1 1 1 1 1 1 ...
##  $ mine.basin               : Factor w/ 1 level "Appalachia Central": 1 1 1 1 1 1 1 1
##  $ mine.status              : Factor w/ 5 levels "Active","Active, men working, not p
##  $ mine.type                : Factor w/ 1 level "Underground": 1 1 1 1 1 1 1 1 1 1 ..
##  $ company.type             : Factor w/ 3 levels "Contractor","Indepedent Producer Op
##  $ operation.type           : Factor w/ 2 levels "Mine only","Preparation Plant": 1 1
##  $ operating.company        : Factor w/ 57 levels "Apple Jacks Coal Company Inc",..:
```

```
##  $ operating.company.address: Factor w/ 61 levels "1229 Homecreek Rd., Big Rock, VA 2
##  $ production.stons         : Factor w/ 163 levels "0","1,719","103,817",..: 87 118 7
##  $ average.employees        : int  16 17 13 14 94 48 37 101 36 7 ...
##  $ labor.hours              : Factor w/ 168 levels "","1,689","101,738",..: 97 127 84
##  $ ARC                      : int  1 1 1 1 1 1 1 1 1 1 ...
```

```
colnames(harlan.raw)
```

```
##  [1] "year"                     "mine.name"
##  [3] "mine.state"               "countystr"
##  [5] "mine.basin"               "mine.status"
##  [7] "mine.type"                "company.type"
##  [9] "operation.type"           "operating.company"
## [11] "operating.company.address" "production.stons"
## [13] "average.employees"         "labor.hours"
## [15] "ARC"
```

```
colnames(dickenson.raw)
```

```
##  [1] "year"                     "mine.name"
##  [3] "mine.state"               "countystr"
##  [5] "mine.basin"               "mine.status"
##  [7] "mine.type"                "company.type"
##  [9] "operation.type"           "operating.company"
## [11] "operating.company.address" "production.stons"
## [13] "average.employees"         "labor.hours"
## [15] "ARC"
```

```
summary(harlan.raw)
```

```
##       year                mine.name             mine.state   countystr
##  Min.   :2000   Mine #1        : 24   Kentucky (East):509   Harlan:509
##  1st Qu.:2003   Prep Plant     : 21
##  Median :2006   Mine No 1      : 19
##  Mean   :2006   Mine No. 1     : 11
##  3rd Qu.:2008   No 1 Plant     : 11
##  Max.   :2011   Darby Fork No 1: 10
##                 (Other)        :413
##               mine.basin                                    mine.status
##  Appalachia Central:509   Active                                   :420
##                           Active, men not working, not producing:  1
##                           Active, men working, not producing     : 34
##                           Mine closed by MSHA                      : 21
##                           Permanently abandoned                    : 27
##                           Temporarily closed                       :  6
##
##        mine.type                        company.type           operation.type
```

```
##  Underground:509   Contractor                    : 92   Mine only        :399
##                    Indepedent Producer Operator:324   Preparation Plant:110
##                    Operating Subsidiary         : 93
##
##
##
##
##                        operating.company
##  Harlan Cumberland Coal Company: 46
##  Lone Mountain Processing Inc  : 31
##  Manalapan Mining Co., Inc.    : 29
##  Manalapan Mining Company Inc  : 28
##  Liggett Mining Llc            : 23
##  Rex Coal Company, Inc.        : 14
##  (Other)                       :338
##                            operating.company.address  production.stons
##  P.O. Box 269, Grays Knob, KY 40829    : 42          0        :110
##  P.O. Box 527, Benham, KY 40807        : 35          1,013,140:  1
##  8174 E Hwy 72, Pathfork, KY 40863     : 31          1,046,339:  1
##  P.O. Box 838, Middlesboro, KY 40965   : 28          1,047,698:  1
##  P.O. Box 1226, Norton, VA 24273       : 25          1,122,057:  1
##  General Delivery, Grays Knob, KY 40829: 22          1,186,166:  1
##  (Other)                                :326          (Other)  :394
##  average.employees  labor.hours       ARC
##  Min.   :  1.00           : 40   Min.   :1
##  1st Qu.: 15.00    23,520 :  3   1st Qu.:1
##  Median : 25.00    1,211  :  1   Median :1
##  Mean   : 34.83    1,692  :  1   Mean   :1
##  3rd Qu.: 40.00    1,702  :  1   3rd Qu.:1
##  Max.   :166.00    1,705  :  1   Max.   :1
##  NA's   :40        (Other):462
```

```
summary(dickenson.raw)
```

```
##       year                       mine.name      mine.state      countystr
##  Min.   :2000   Cherokee Mine         : 12   Virginia:194   Dickenson:194
##  1st Qu.:2001   Mc Clure River  Plant: 10
##  Median :2004   Nme                   :  8
##  Mean   :2005   No 8                  :  7
##  3rd Qu.:2007   No. 3 Mine            :  7
##  Max.   :2011   No. 4                 :  7
##                 (Other)               :143
##                mine.basin                                  mine.status
##  Appalachia Central:194   Active                                 :151
##                           Active, men working, not producing: 14
```

```
##                           Mine closed by MSHA               :  4
##                           Permanently abandoned             : 21
##                           Temporarily closed                :  4
##
##
##       mine.type                            company.type           operation.type
##  Underground:194   Contractor                    :50    Mine only        :162
##                    Indepedent Producer Operator:66    Preparation Plant: 32
##                    Operating Subsidiary         :78
##
##
##
##
##                        operating.company
##  DickensonRussell Coal Company : 19
##  Clinchfield Coal Company      : 12
##  DickensonRussell Coal Co Llc  :  9
##  L & J Equipment Company       :  9
##  Apple Jacks Coal Company, Inc.:  7
##  DickensonRussell Coal Co., Ll :  7
##  (Other)                       :131
##                           operating.company.address production.stons
##  P.O. Box 1426, Grundy, VA 24614      : 15         0      : 32
##  5703 Crutchfield Dr, Norton, VA 24273: 11         1,719  :  1
##  P.O. Box 1025, Grundy, VA 24614      :  9         103,817:  1
##  P.O. Box 148, Vansant, VA 24656      :  8         104,108:  1
##  P.O. Box 458, Big Rock, VA 24603     :  8         104,515:  1
##  Rt 2, Box 73, Cleveland, VA 24225    :  8         109,076:  1
##  (Other)                              :135         (Other):157
##  average.employees  labor.hours       ARC
##  Min.   :  1.00           : 26   Min.   :1
##  1st Qu.: 12.00    79,091 :  2   1st Qu.:1
##  Median : 16.00    1,689  :  1   Median :1
##  Mean   : 30.11    101,738:  1   Mean   :1
##  3rd Qu.: 40.00    104,181:  1   3rd Qu.:1
##  Max.   :117.00    107,901:  1   Max.   :1
##  NA's   :26        (Other):162
```

3 Exploratory Analysis

3.1 Visual Data Exploration of Harlan and Dickenson Data

The relationship of mines to labor hours, employment, and production across the coal sector from 2000 to 2011 does not seem to be a consistent trend, nor is it easy to visualize (Figures 1, 3, and 5). There are too many mines to determine any sort of relationship; furthermore, whether a mine is from a boom-bust or bust-bust county is indeterminable in these visualizations (Figures 1, 3., and 5). However, mine data for employees related to production shows the most promise for a significant correlation at the mine level; however, since this project focuses on coal counties, this correlation is not especially relevant to the research question (Figure 7).

On the other hand, the exploratory graphs indicate that there are strong trends in the county-wide data (Figures 2, 4, and 6). Labor hours are greater than in Harlan than Dickenson on the whole, but do not seem to display a steady trend through the years (Figure 2). What's interesting is Harlan appears to have substaintal fluctuations in production through the years, decreasing quickly in 2002, while Dickenson steadily declined (Figure 6). Employment seems to follow similar trends for both counties, but not as drastically (Figure 4). Like the mine level data, both counties seem to have a strong correlation for employees with total production (Figure 8).

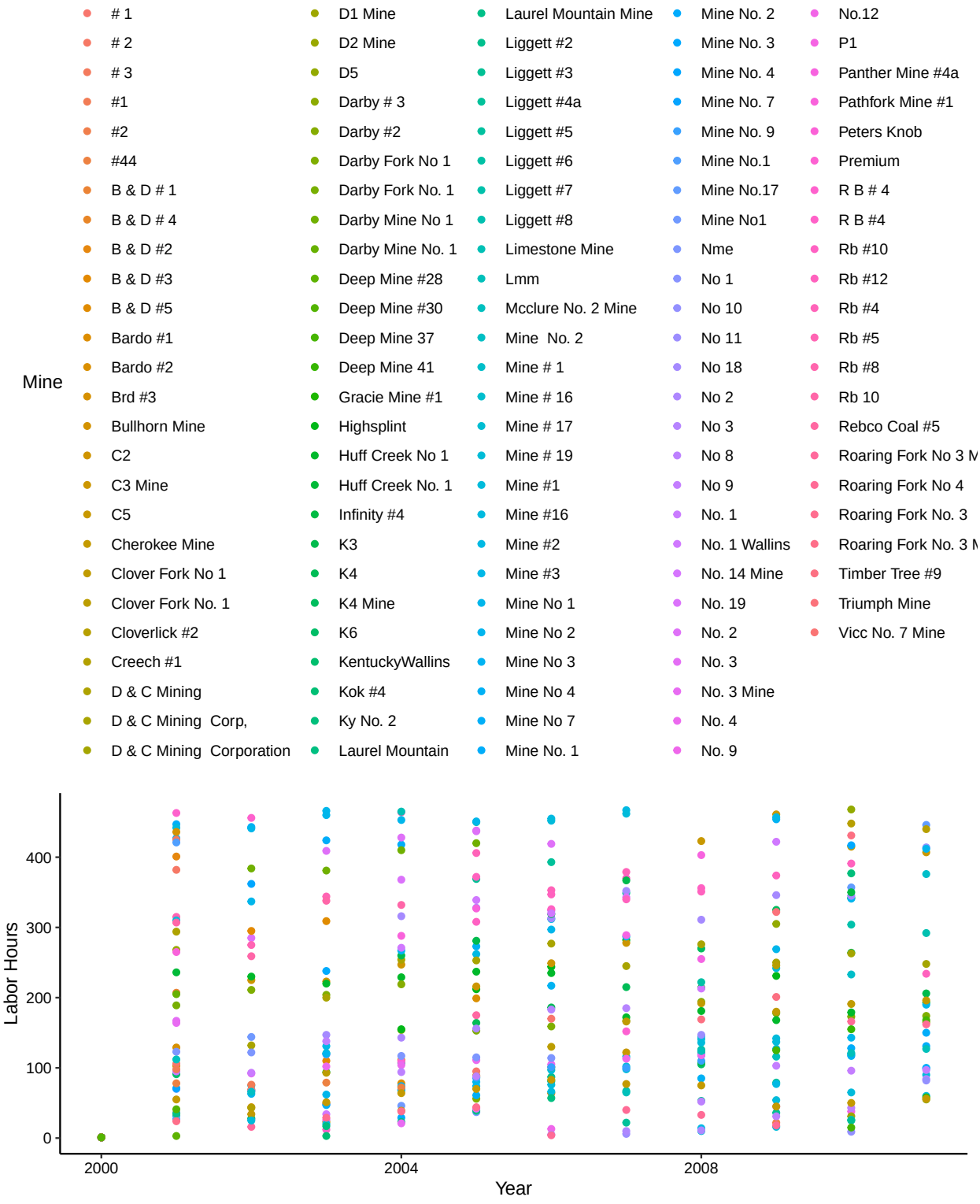

Figure 1: Total Annual Labor Hours for Mining in Harlan and Dickenson Counties by Mine

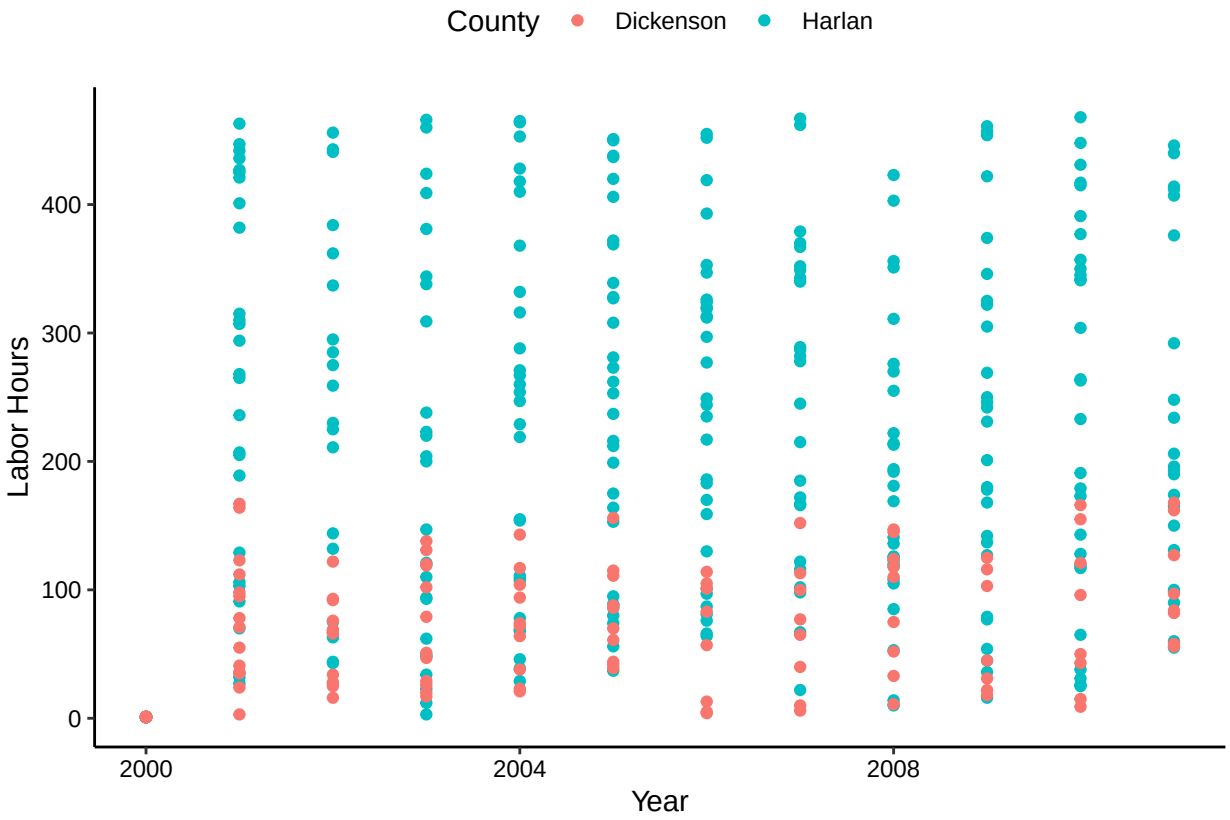

Figure 2: Total Annual Labor Hours for Mining in Harlan and Dickenson Counties by County

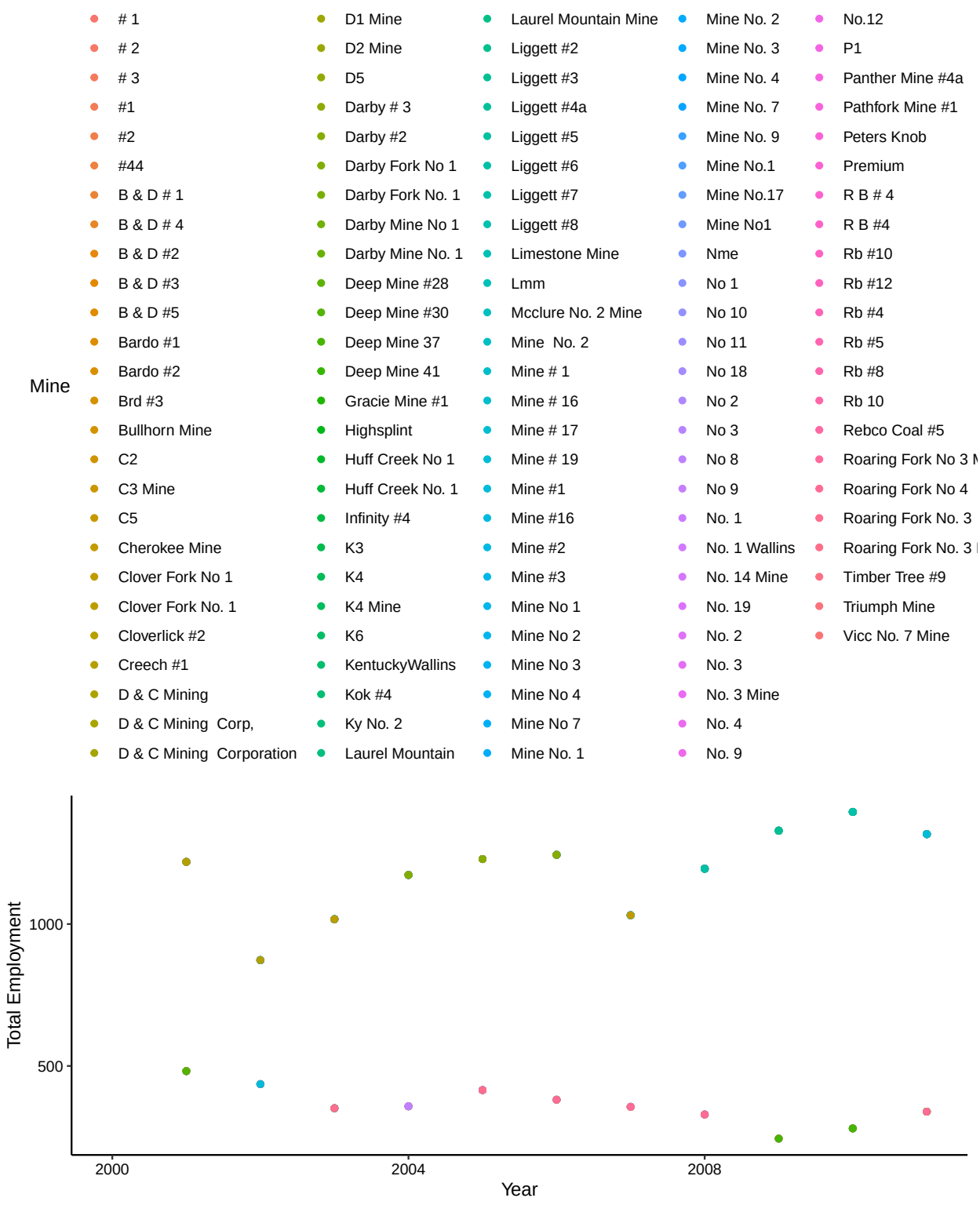

Figure 3: Total Annual Coal Employment in Harlan and Dickenson Counties by Mine

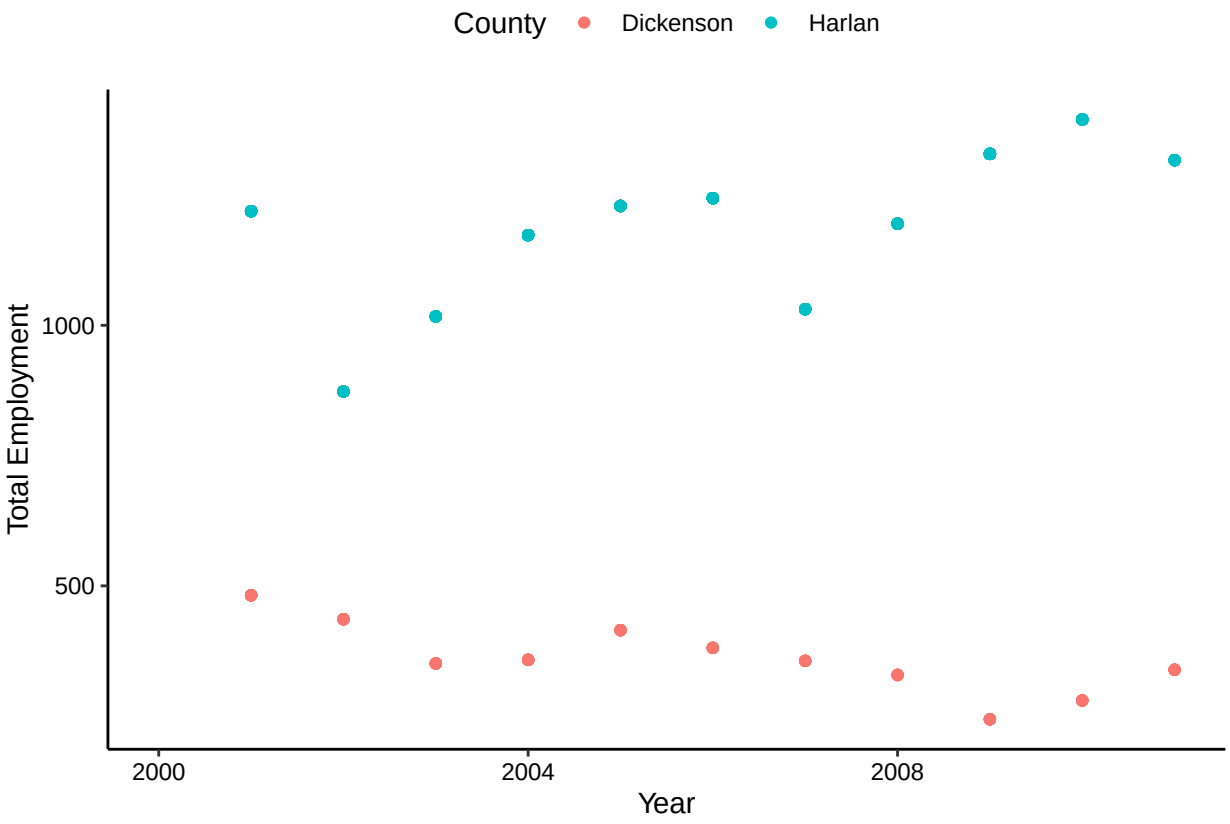

Figure 4: Total Annual Coal Employment in Harlan and Dickenson Counties by County

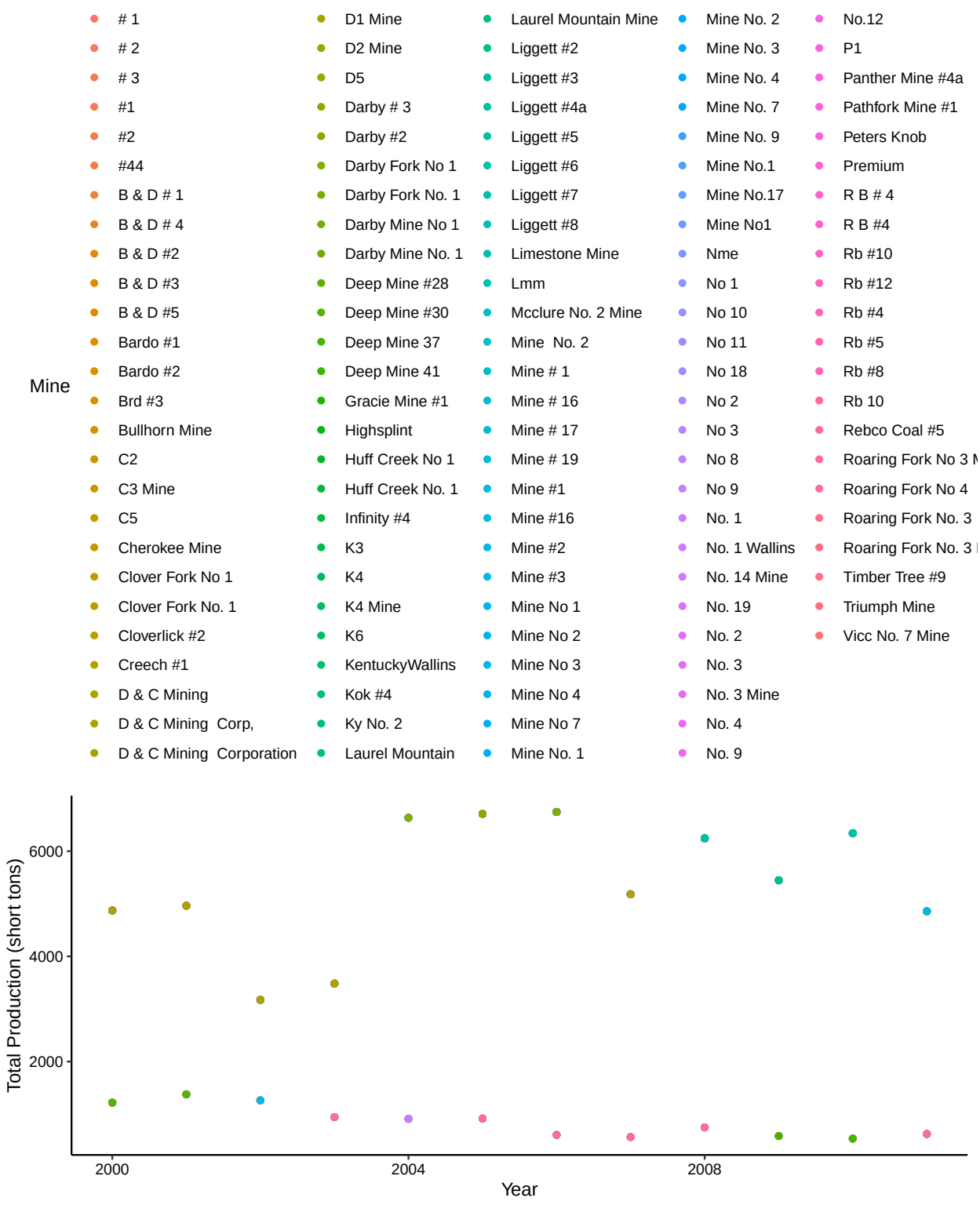

Figure 5: Total Annual Coal Production in Harlan and Dickenson Counties by Mine

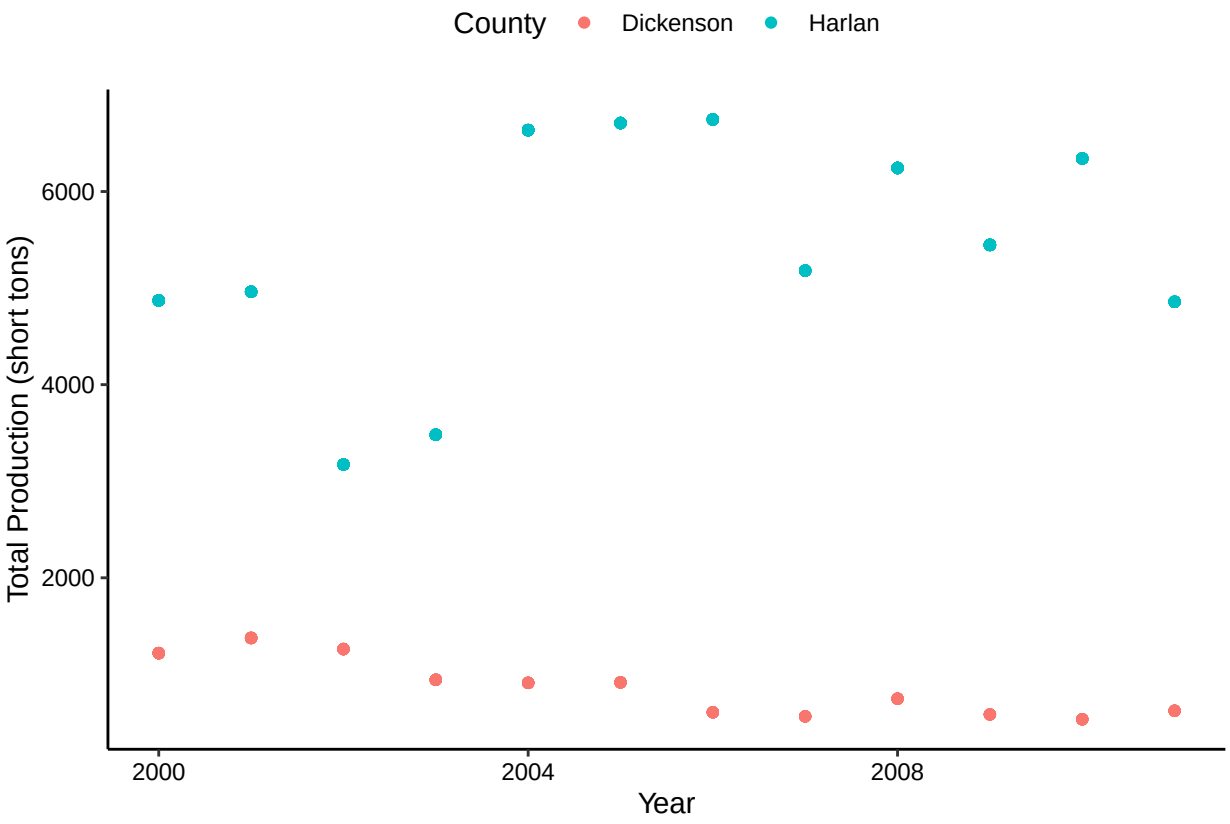

Figure 6: Total Annual Coal Production in Harlan and Dickenson Counties by County

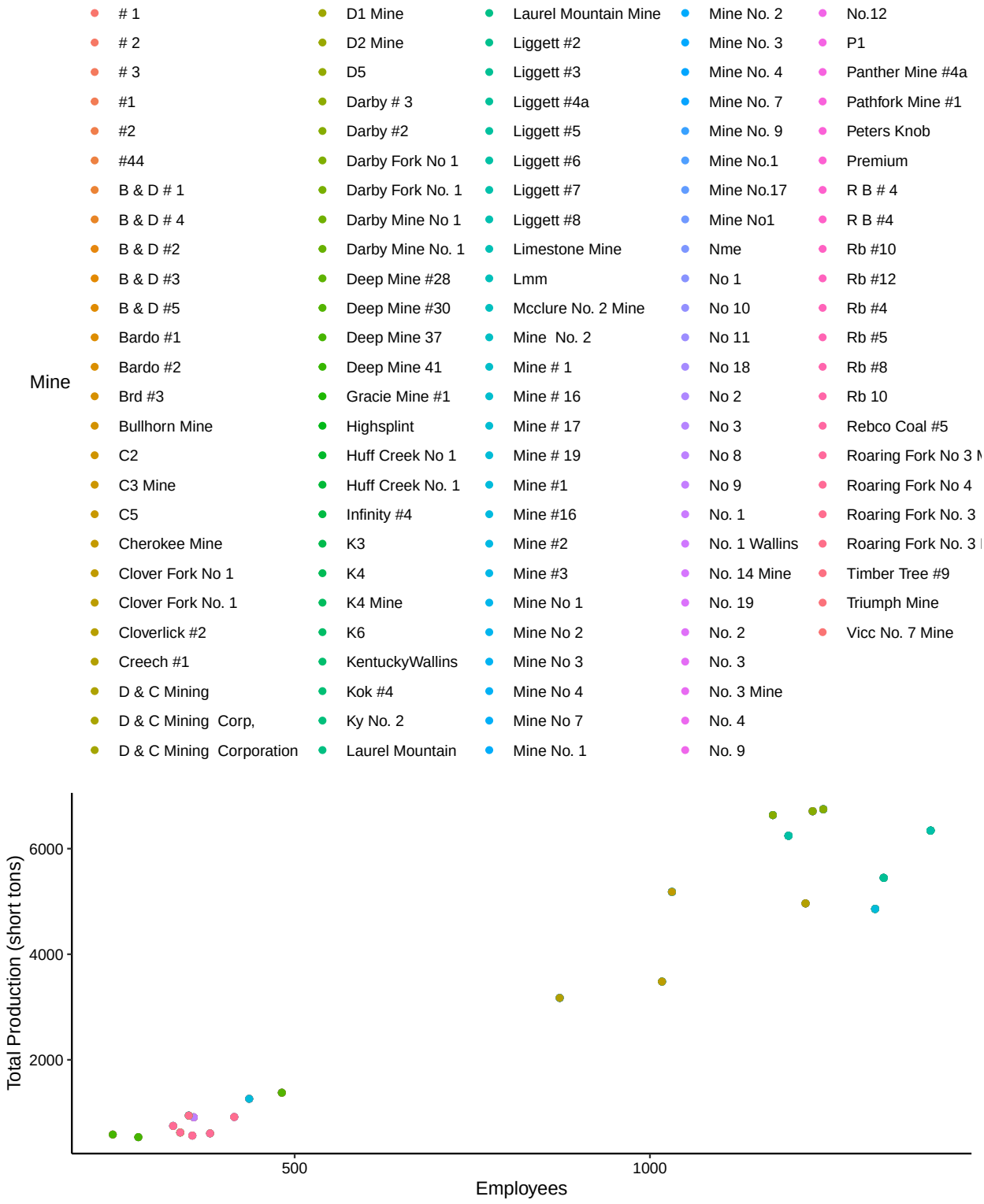

Figure 7: Total Annual Coal Production versus Total Employment in Harlan and Dickenson Counties by Mine

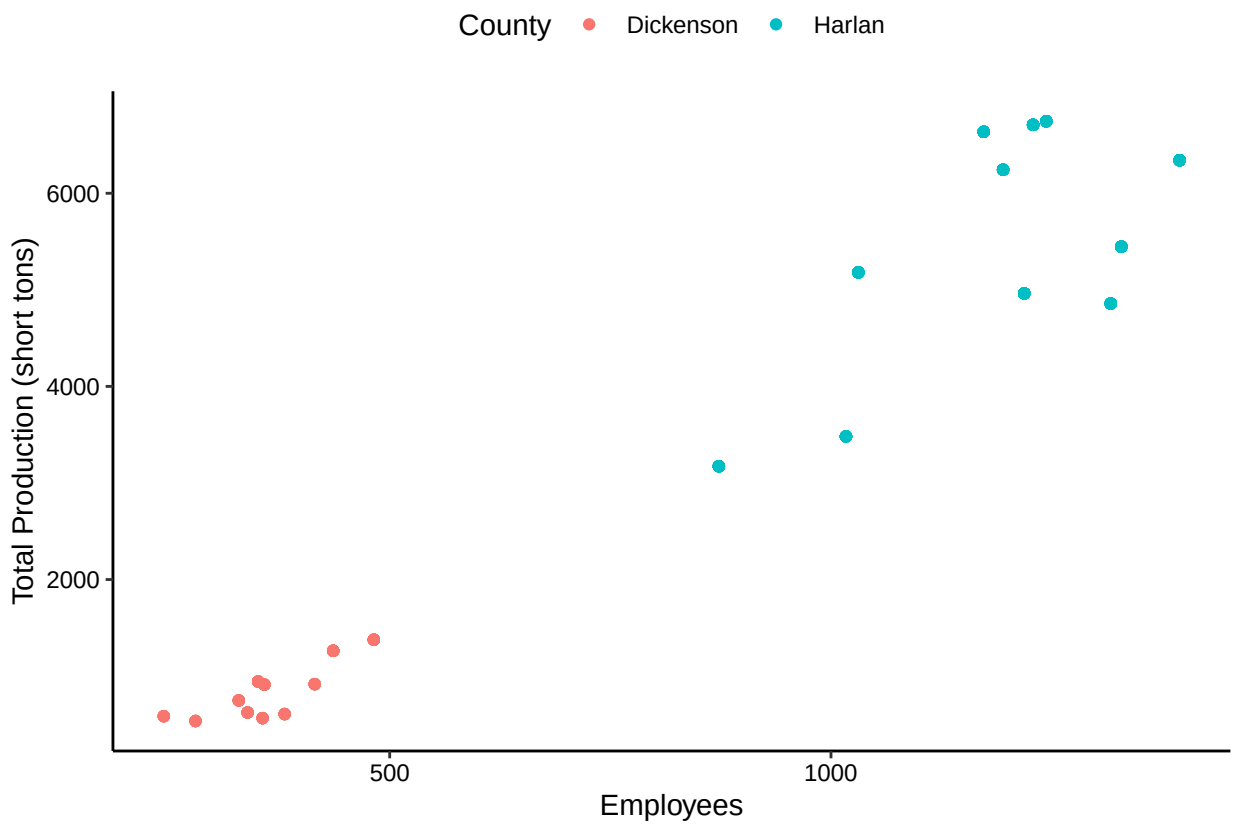

Figure 8: Total Annual Coal Production versus Total Employment in Harlan and Dickenson Counties by County

4 Analysis

4.1 What is the relationship between coal production and number of employees in coal mines in Harlan and Dickenson Counties?

4.2 How does coal production differ between all mines in Harlan County compared to Dickenson County?

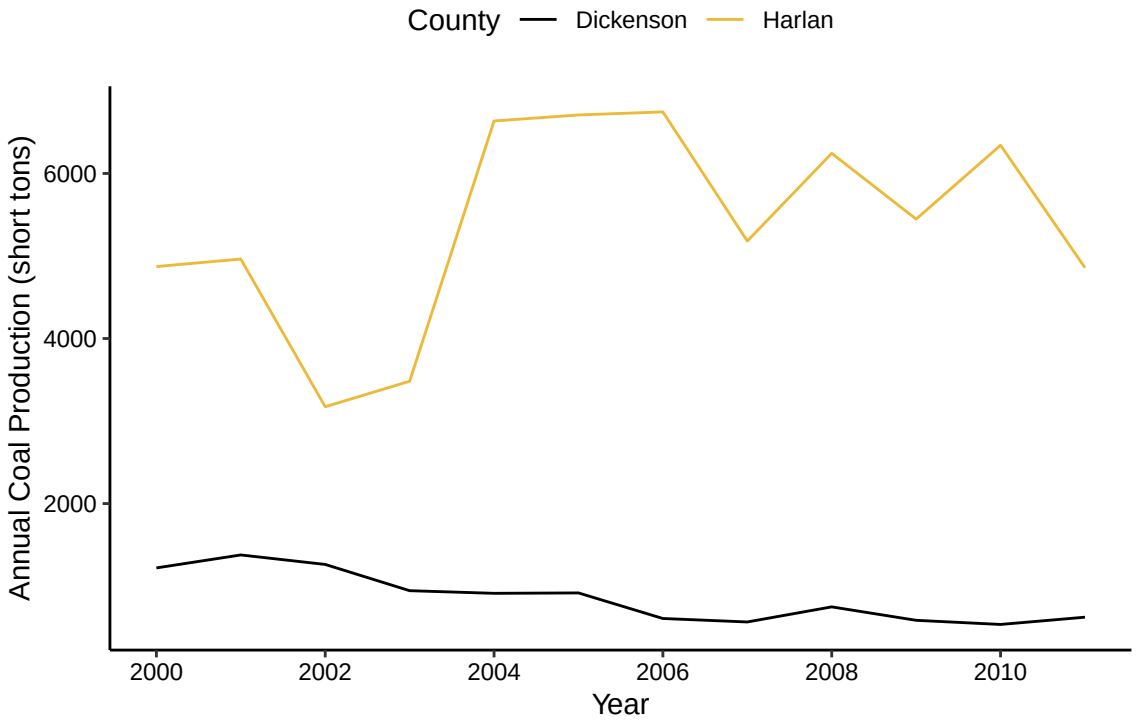

Figure 9: Total Annual Coal Production in Harlan and Dickenson Counties

4.3 How does the number of employees in the mines of Harlan County compare to Dickenson County?

4.4 Is the total number of employees for all mines in a boom-bust and a bust-bust county in the years 2000 to 2011 a significant predictor of annual coal production for all mines?

Harlan County, Kentucky produced a much greater amount of coal and employed more people than Dickenson County, Virginia throughout the early twenty-first century (Figure 9 and Figure 10). This divergence is to be expected with Harlan being boom-bust and Dickenson being bust-bust for coal production 2000 to 2011.

However, Harlan County experienced a variable production amount from year to year, which is consistent with historical coal production in the twentieth century (EIA, 2016). Dickenson

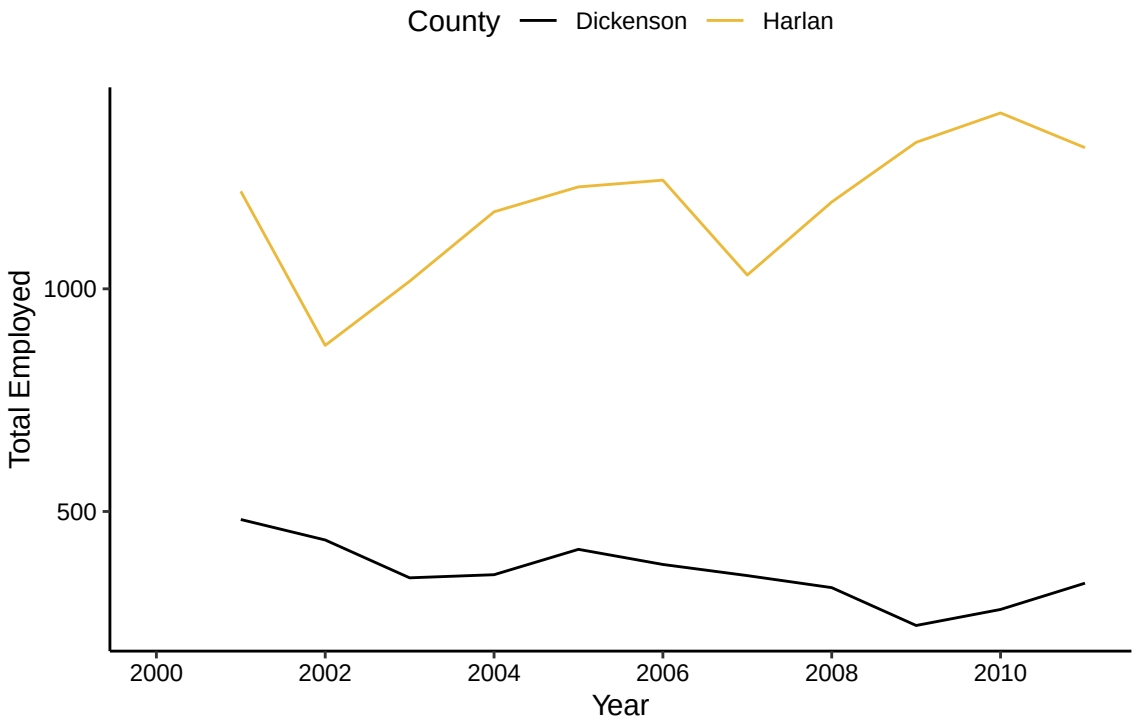

Figure 10: Total Number of Employees in Harlan and Dickenson County Mines, 2000 to 2011

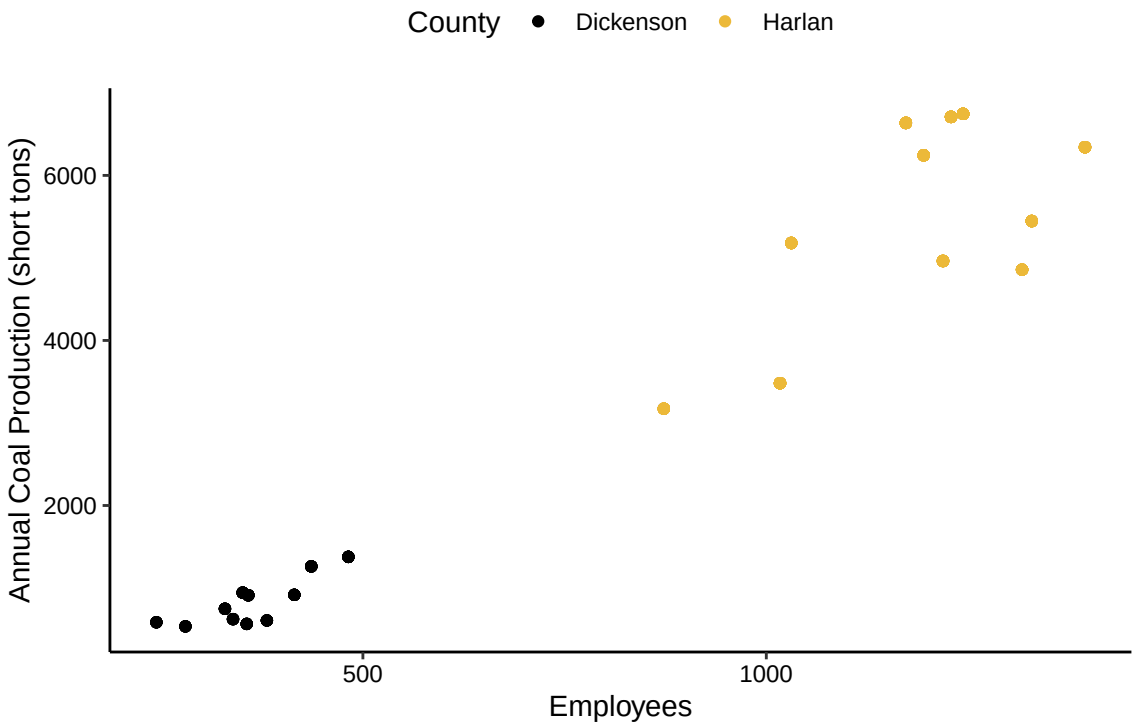

Figure 11: Annual Coal Production Compared to Total Employees in Harlan and Dickenson Counties, 2000 to 2011

County, on the other hand, experienced a steady decline in coal production since 2000 (Figure 9).

Surprisingly, Dickenson also experienced a small increase in coal employment in 2009 even though there is no coal production increase during that same time period (Figure 9 and Figure 10). As such, it becomes essential to question if mine production is an accurate predictor of mine employment. If so, coal production could be correlated to migration from an area based on employment.

Although Harlan had more employees and coal production overall than Dickenson, coal production significantly impacts the number of employees in mining for both counties (Figure 11; ANOVA; Harlan: df = 288, F = 197.2, p-value < 0.001; Dickenson: df = 105, F = 276.6, p-value < 0.001).

5 Summary and Conclusions

Although Harlan and Dickenson counties are both "coal counties" within the scope of this research, Harlan experienced a boom in coal production in 2000 but a bust in 2010. Dickenson steadidly decreased in that same time period. However, Harlan had variable coal production and employment, exhibiting smaller boom-bust cycles within the decade.

Most importantly, both counties' coal production signficantly correlated with employment in coal mines. Intuitively, this makes sense, as lower production means less money means less available jobs.

Finally, the correlation between production and employment could be important for migratory trends in Appalachia. Based on the data in this research, their are less emplyment opportunities in coal mine with less coal production. As such, residents of Appalachia may leave the region in search of jobs as the energy sector slowly transitions from coal-fired power plants to forms of cleaner energy.

6 References

EIA. (2020). **Annual Survey of Coal Production and Preparation; Form EIA-7A** https://www.eia.gov/survey/form/eia_7a/form.pdf

EIA. (2016). **Quarterly coal production since the early 1980s**. https://www.eia.gov/todayinenergy/detail.php?id=26612