

# Criminal Mines: Evaluating the effects of the coal mining industry on incarceration and drug-related crime

## Methodology

Robert Betancourt<sup>\*</sup>, Connor Bulgrin<sup>\*</sup>, Jenny Duan<sup>†</sup>, Nicholas Skelley<sup>\*</sup>, and  
Adeline Sutton<sup>\*</sup>

<sup>\*</sup>Department of Economics, Cornell University

<sup>†</sup>Jeb E. Brooks School of Public Policy, Cornell University

Last updated: November 23, 2025

## 1 Background

Between 2001 and 2023, aggregate coal productivity in the US has gone from 1.13 billion short tons to 578 million, a 48.75% decrease (U.S. Energy Information Administration, [2023](#)). In the same period, natural gas growth withdrawals have increased from 24.5 million (million cubic feet) to 45.6 million (U.S. Energy Information, [2025](#)). Prior research has explored the effect of the coal decline on social effects. Blonz et al. ([2023](#)) find decreased credit scores and increased use of credit, even when accounting for individuals who lost mining-related jobs. Welch and Murray ([2020](#)) study the impact of coal activity on local revenue, finding a decrease in local revenue for education per student. Naturally, we expect more poorly funded counties to have worse justice outcomes. A minority of the literature finds the opposite conclusion—for example, Nason et al. ([2025](#)) find some evidence that long-term local reduction in poverty was due to outmigration, wage increases and increased educational investment. We contribute to the literature by studying the relationship between coal deindustrialization and crime at the county level using a continuous treatment difference-in-differences approach based on Callaway et al. ([2024a](#)).

Our broad question of interest is “How does deindustrialization impact crime outcomes?” More specifically, we test whether a decline in a county’s share belonging to coal causes an increase in various crime statistics, including violent, property, and drug-related crime.

## 1.1 De-industrialization and exposure shock

Most other papers that look at declines in industrialization or manufacturing that use exposure shocks use exposure as a continuous variable. Howard and Manuelli (2024) and Glaeser and Saiz (2003) both use baseline manufacturing employment shares (the peak of U.S manufacturing employment) as a measure for manufacturing exposure in different years. Similarly, Gagliardi et al. (2023) looks at the effect of de-industrialization on employment consequences in cities across six countries and uses the city’s share of manufacturing employment in the year of its country’s manufacturing peak as treatment for subsequent changes in total employment. Former manufacturing hubs are defined as cities with an initial manufacturing employment share in the top tercile of their country’s distribution, while middle and bottom terciles are used for comparison in their 2SLS regression using driving distance to historical colleges and universities as an instrument. Autor et al. (2013) analyze the effects of rising Chinese import competition between 1990 and 2007 on US local labor markets, exploiting cross-market variation in import exposure stemming from initial differences in industry specialization. They define exposure as local industry employment shares in time  $t$  and exploit variation in local labor market exposure to Chinese import competition.

## 2 Methodology

The challenge of answering whether coal deindustrialization (here, decline in production) *causes* a change in crime outcomes is the possibility of endogeneity. That is, there may be other factors correlated with production that also impact crime, or reverse causality whereby firms actually exit because of crime levels in the county. We argue that the period immediately following a peak in coal production marks the beginning of deindustrialization and is exogenous to crime outcomes.

We use an event study to estimate the impacts of production decline after a county’s largest peak in coal production. The year immediately following the peak is the event, so event time is measured in years after the initial decline. See Figure 2 for an aggregated visualization of coal production peaks. We measure dynamic treatment effects in order to see how the decline in coal production impacts a county over time. That is, in each year after the initial decline, how is that initial decline still impacting a county?

## 2.1 Continuous treatment event studies

The standard binary treatment event study can be extended to the continuous treatment case. The following adapts the description from Callaway et al. (2024a) to our purpose.

### 2.1.1 Potential outcomes and treatment effects

In a standard event study, the treatment is an event; a unit is untreated before the event and treated after. The potential outcomes are  $Y_{kt}(0)$  for individual unit  $k$ ’s outcome in time  $t$  if  $k$  was untreated and  $Y_{kt}(1)$  if  $i$  were treated. We can extend this to the continuous case where  $Y_{kt}(d)$  is the outcome of  $k$  in time  $t$  had the individual received “dose”  $d$ . In our context,  $Y_{kt}(d)$  would be county  $k$ ’s crime outcome in year  $t$  if they had an initial decline in coal production from their peak of slope  $d$ .

While in a standard event study the treatment effect is the difference between the treated and untreated potential outcomes, in the continuous case there are two types of causal effects: the level treatment effect and the causal response. The level treatment effect of dose  $d$  is the difference in the potential outcome given the individual received dosage  $d$  and the potential outcome given the individual was untreated:  $Y_{kt}(d) - Y_{kt}(0)$ . The causal response of dose  $d$  is the slope of the level treatment effect at  $d$  with respect to the dosage:  $Y'_{kt}(d)$ ; this measures the effect of increasing the dosage incrementally. In our context, the level treatment effect at  $d$  would be the difference in the potential outcomes of a county when they had a decline of  $d$  compared to if they had no decline, while the causal response at  $d$  would be how the outcome of county  $k$  would change if they had a decline of  $d + \varepsilon$  compared to the actual dose  $d$ , which gets at how impactful having higher rates of

decline is.

### 2.1.2 Parameters of interest

The two parameters of interest then are the average level treatment effects and the average causal responses. We consider the case where once an individual is treated with a particular dose, they remain treated with that dose thereafter. Let  $g$  denote the time the individual becomes treated. The average treatment effect (ATE) of dosage  $d$  is

$$ATE(g, t, d) = \mathbb{E}[Y_{kt}(g, d) - Y_{kt}(0) | G = g],$$

and the average causal response is

$$ACR(g, t, d) = \frac{\partial ATE(g, t, d)}{\partial d} = \frac{\partial Y_{ktg}}{\partial d}.$$

Under assumptions discussed in 2.3, the ATE parameter is identified by  $\beta_j$  in the estimating equation below.

## 2.2 Specification

### Estimating equation

$$y_{kt} = \alpha + \sum_{j \neq -1} (\beta_j \cdot \delta_{kt} \cdot \mathbb{1}[j = t]) + \theta_t + \phi_k + \varepsilon_{kt}$$

- $y_{kt}$  is an outcome in county  $k$  in calendar year  $t$ .
- For event time  $j$ ,  $j \equiv -1$  in the year of the county  $k$ 's coal production global (within time span of interest, omitting counties with edge maxima) maximum.
- $\delta_{kt}$  is the treatment intensity/dose—the *percent change* in coal production in county  $k$  in time  $t$  relative to event time  $-1$ . This allows us to investigate heterogeneity in treatment effects by how quickly production declines.

- $\theta_t$  and  $\phi_k$  are calendar year and county fixed effects.

The dynamic treatment effects in this event study are  $\beta_j$ ; this is the effect of a county's coal production decline on crime outcomes  $j$  periods after the peak of production.

## 2.3 Assumptions

As this is an event study with continuous treatment intensities, it is subject to similar but stronger assumptions in order to interpret the treatment effect as causal (that is, the decline in the coal industry drives changes in crime). These are:

1. **No anticipation** This would imply that agents did not foresee a decline in coal production and changed their behavior as a result.
2. **Strong parallel trends** Strong parallel trends “says that the average evolution of outcomes for the entire population if all experienced dose  $d$  is equal to the path of outcomes that dose group  $d$  actually experienced” (Callaway et al., 2024a). In comparison, the parallel trends assumption in an event study says that the evolution of outcomes of treated and untreated individuals would be the same in the absence of treatment. Under the assumption of parallel trends but not *strong* parallel trends, the event study identifies the average treatment on the treated (ATT) and is subject to selection concerns. According to Callaway et al. (2024b), these selection concerns preclude causal interpretation because observed low-dose outcomes are not valid counterfactuals for what high-dose individuals would experience at the low dose.

Based on the event study figures in 3a and 3b, the lack of a significant difference from the rates in each year leading up to the peak suggests that counties pre-decline have similar trends in their crime outcomes. This lends evidence to support the above assumptions.<sup>1</sup> If there had been anticipation, we would expect to see a jump in coefficients on outcomes right before event time 0.

---

<sup>1</sup>We stumbled upon this recent work on continuous difference-in-differences methods and assumptions relatively recently and are still in the process of identifying the best ways to go about justifying the strong parallel trends assumption, a task made somewhat more challenging given the recency of these developments in the methodology and the sparsity of empirical work applying these new methods

If there had not been parallel trends, we would expect to see significant coefficients before 0 (this would mean that there is a differential effect between treated and untreated groups even before treatment).

### 3 Data

Our data come from three primary sources. From CJARS, we obtain criminal justice outcomes aggregated by county, cohort year, sex, race, and binned age group. Table 1 provides detailed variable definitions for the CJARS measures used in our analyses. To capture more detailed demographic and economic characteristics at the county level, we supplement with data from the American Community Survey (ACS) on age, sex, race, educational attainment, and wage and salary income. Finally, data on coal production by county come from detailed, mine-level data from the U.S. Energy Information Administration (EIA) and Mine Safety and Health Administration (MSHA). We aggregate coal data to the county-year level and obtain county-level raw coal production quantities for each year from 1983 to 2023. Our primary outcome variables are felony and misdemeanor rates, measured as the number of criminal charges per 100,000 county residents.

### 4 Results

Descriptively, we include Figure 1a to provide the number of counties with peak coal production for each calendar year. For the purposes of our event study, and because CJARS observations are sometimes missing for smaller counties, we also include Figure 1b, which describes counts of peak coal production for each calendar year, restricted to counties whose peak production years are strictly in the interior of their CJARS observations. We observe that within the full analysis sample, peak coal production appears to occur disproportionately before 2010. The years of peak coal production, however, are dispersed broadly over the available range, lending credibility to the assumption that there are no exogenous, contemporaneous shocks that may confound our interpretation of any effect observed in the event study.

Figure 2a shows the average county-level  $z$ -scores without weights. Coal production varies

Table 1: CJARS Variable Definitions

| Variable       | Description   |
|----------------|---|
| fips           | County FIPS code  |
| cohort_year    | Year of criminal justice event (charge, disposition, prison entry/exit, probation/parole start/end) |
| sex            | Sex   |
| race           | Race  |
| age_group      | Age group   |
| off_type       | Offense type  |
| repeat_contact | Repeat contact indicator  |
| fe_rate        | Felony rate per 100,000 population  |
| N_fe_rate      | Resident population for felony statistic  |
| mi_rate        | Misdemeanor rate per 100,000 population   |
| N_mi_rate      | Resident population for misdemeanor statistic   |
| inc_rate       | Incarceration rate per 100,000 population   |
| N_inc_rate     | Resident population for incarceration statistic   |
| par_rate       | Parole rate per 100,000 population  |
| N_par_rate     | Resident population for parole statistic  |
| pro_rate       | Probation rate per 100,000 population   |
| N_pro_rate     | Resident population for probation statistic   |

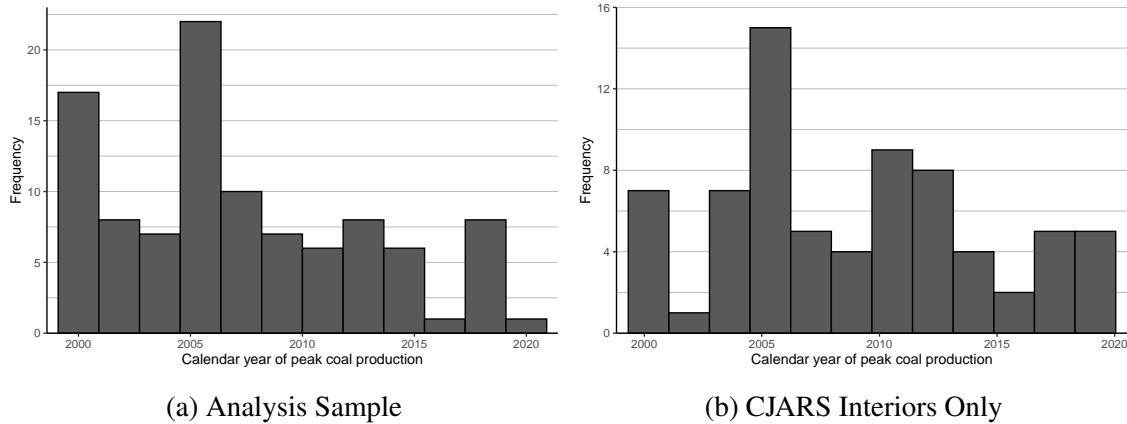


Figure 1: Distribution of Calendar Years of Peak Coal Production

*Notes:* Histograms showing the distributions of calendar years in which counties achieved their peak coal production. The interiors-only figure, Figure 1b, shows the distribution for the analysis sample of counties, restricted further to exclude counties whose peak production years are their first year observed in the CJARS data.

widely across counties, so we normalize coal production within each county using the within-county  $z$ -score of production quantity. An event time of 0 indicates the first year after each county's coal production peak. As expected, this figure indicates that average coal production increases up to event time  $-1$  and then declines. Figure 2b depicts the same figure where each county is weighted by the raw quantity of coal produced in the county at event time  $-1$ . Notably, the average rate at which production increases before reaching peak production is visually similar to the rate of decline after the peak. This observation raises the concern that if coal production affects crime outcomes symmetrically for increases and decreases in production, then we may either observe significant pre-event trends or no post-event trends, the latter of which occurs.

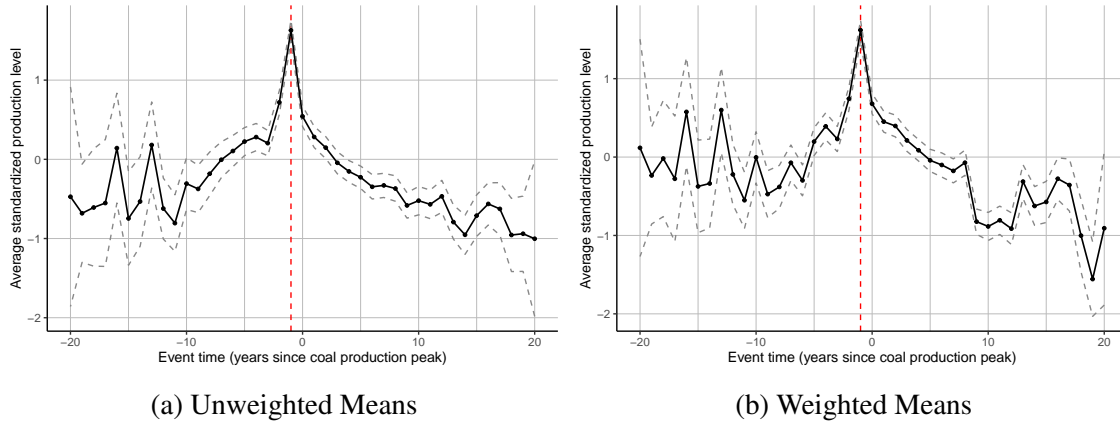


Figure 2: Average County-Level  $z$ -Scores

*Notes:* Average county-level coal production quantity by year, expressed as within-county standard deviations from the county mean level of production (within-county  $z$ -scores). Dashed lines show unweighted 95% confidence intervals using year-level standard errors. In the weighted graph, Figure 2b, annual means are weighted at the county level by the raw quantity of coal produced in the county at event time  $-1$ , the year of peak coal production.

Figures 3a and 3b depict our event studies of interest. They present the average change in the annual charge rate (number of charges per 100,000 residents) relative to the year in which coal production was the highest. In both of these, the lack of a significant difference from the rates in each year before event time 0 is evidence that pre-trends hold. Similarly, the lack of a jump in average change in the years leading up to event time 0 is evidence that our no anticipation assumption holds.

Figure 3a suggests that the number of years since the peak in coal production did not have



a significant effect on the rate of felony charges. At first glance, Figure 3b indicates that the number of years since the peak may have had a slight increase on the rate of misdemeanor charges. However, for many of these years, we cannot reject the null at a 95% confidence interval. We are also skeptical of the validity of the parallel trends assumptions, as visual inspection suggests that the misdemeanor rate may have been increasing prior to the event time and merely continued afterwards.

These findings are inconsistent with much of the literature. For instance, Welch and Murray (2020) find that coal deindustrialization negatively impacted local education revenue, which makes our null finding unexpected. Similarly, Schept (2022) argues that the pattern of extraction, crisis, and profiteering contributed to rising incarceration rates in Central Appalachia. Our results suggest that the deindustrialization of coal in the US did not substantially impact felony and misdemeanor crime rates. However, these results should be interpreted with caution. Among other concerns and opportunities for more robust analysis discussed in Section 5, the relatively small sample of counties we can observe (212 counties in the panel with valid felony rates, and 207 counties in the same panel for misdemeanors) leaves us unable to reliably detect small effects.

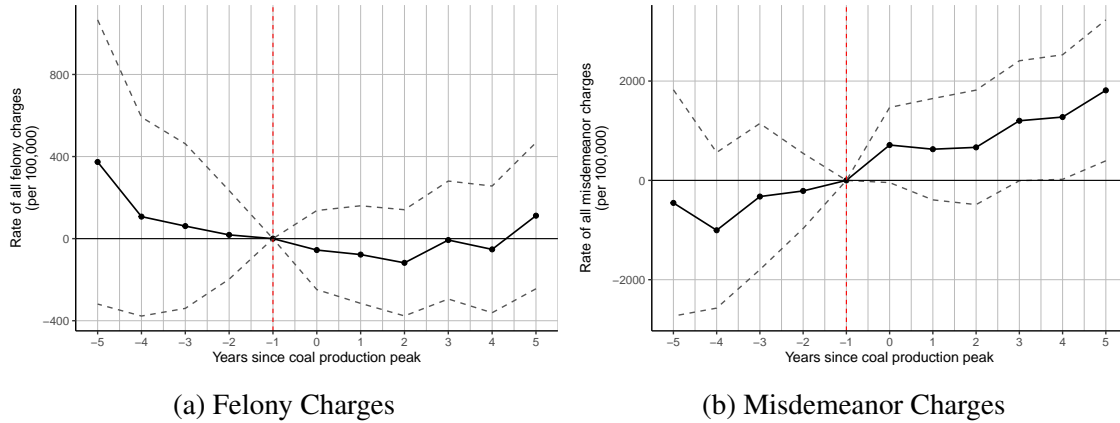


Figure 3: Event Studies of Crime Rates

*Notes:* Point estimates of *ATE* parameters under our main, continuous-treatment event study specification that includes fixed effects for county and calendar year. Point estimates represent average change in the annual rate of charges (number of charges per 100,000 county residents) relative to the year in which the county's coal production was highest. Dashed lines show 95% confidence intervals using standard errors clustered by county. Event time  $-1$  is the year in which coal production was highest in the county. Estimates are computed for the minimal balanced sample of counties who are observed for at least five years on each side of their coal production peak ( $N = 212$  for felonies and  $N = 207$  for misdemeanors).

## 5 Discussion

Although we find null results that suggest that the number of years since coal production peak (deindustrialization) did not have a significant impact on felony rates at the county level, our identification strategy faces several potential threats to validity, and further robustness checks are needed to confirm our findings.

We first note that our continuous-treatment event study design assumes that the dose level remains constant after the event: that a unit is treated at dose  $d$  and maintains that dose thereafter. In our setting, we are then identifying only the effect of the steepness of the initial decline in coal production, and not the trend in subsequent years. This restriction limits our interpretation, as it might be that after the initial decline, two counties with the same initial decline diverge such that one drops more sharply than the other (faster deindustrialization). Under the current specification, these two counties are assumed to have the same dose, even though their crime outcomes may change heterogeneously due to their later divergence. We plan to investigate dynamic dose strategies to flesh this possibility out, though this appears to be beyond the current frontier in the difference-in-differences literature.

A key concern is that coal production and employment may decline due to broader macroeconomic conditions rather than industry-specific shocks. For instance, coal production may fall during recessions due to weakening global economic conditions and reduced energy demand, while criminal justice outcomes may also respond to these same macroeconomic factors through unemployment and economic distress. Specifically, in our data, we are concerned about the confounding effect of the 2008 financial crisis, as the global recession could plausibly affect both our treatment and outcome variables. The relationship between coal production peaks and crime could reflect common responses to business cycles rather than causal effects of deindustrialization. To address this threat, future robustness checks should include state-by-year fixed effects, which would absorb any state-level policy responses to recessions (such as changes in criminal justice enforcement or social safety net programs) as well as common macroeconomic shocks affecting all counties within a state. This would isolate effects from within-state, cross-county variation independent of state-level trends.

Our use of coal production as the treatment variable, rather than coal employment, presents another limitation. Coal production can increase due to technological improvements in productivity per worker, even as employment declines. Our hypothesized mechanisms connecting deindustrialization to crime outcomes are labor market disruptions, including unemployment and transitions; coal employment would be a more direct measure of the industry shock, albeit one more subject to local endogeneity. U.S. coal employment peaked in the early 1900s, well before our CJARS data coverage, which begins in cohort year 2000. This temporal mismatch means our event study may misidentify the timing of unemployment effects. Counties that reached peak coal production in our data may have experienced decades of declining coal employment prior. Future robustness checks should incorporate coal industry wage bill data, which would better capture employment-specific effects by weighting employment changes by compensation levels. This approach would help isolate the influence of industrial *employment* on crime. Additionally, if we can get access to longer-run crime data, using peak coal employment rather than peak production would provide a more direct test of the effect of deindustrialization on crime outcomes.

## References

- Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6), 2121–2168.
- Blonz, J., Tran, B. R., & Troland, E. E. (2023). *The Canary in the Coal Decline: Appalachian Household Finance and the Transition from Fossil Fuels* (NBER Working Paper No. 31072). <https://doi.org/10.3386/w31072>
- Callaway, B., Goodman-Bacon, A., & Sant’Anna, P. H. C. (2024a). *Difference-in-differences with a Continuous Treatment* (NBER Working Paper No. 32117). Cambridge, MA.
- Callaway, B., Goodman-Bacon, A., & Sant’Anna, P. H. C. (2024b). Event Studies with a Continuous Treatment. *AEA Papers and Proceedings*, 114, 601–605. <https://doi.org/10.1257/pandp.20241047>
- Finlay, K., Mueller-Smith, M., & CJARS Team. (2024). Justice Outcomes Explorer (JOE).
- Gagliardi, L., Moretti, E., & Serafinelli, M. (2023). The rust belt’s silver lining: Manufacturing decline and the rise of higher education. *Working Paper*.
- Glaeser, E. L., & Saiz, A. (2003). The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs*, 47–105.
- Howard, G., & Manuelli, R. (2024). Universities and the great reallocation. *Working Paper*.
- Nason, D., Scavette, A., & Stephens, H. M. (2025). Reversing the Natural Resource Curse: Coal Mine Closures and Poverty in Appalachia.
- Schept, J. N. (2022). *Coal, Cages, Crisis: The Rise of the Prison Economy in Central Appalachia*. NYU Press.
- U.S. Energy Information, A. (2025). U.S. Natural Gas Gross Withdrawals (Million Cubic Feet).
- U.S. Energy Information Administration. (2023). Coal Data Browser.
- Welch, J. G., & Murray, M. N. (2020). The Impact of Coal Activity on Local Revenues for Elementary and Secondary Education in Appalachia. *Public Finance Review*, 51(4), 568–613. <https://doi.org/10.1177/10911421221149998>