"The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust"

Report and Replication - Grant Holtes 831099

Introduction

The purpose of the paper *"The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust"* by Black, Daniel and Sanders is to investigate the impact of economic incentives on disability program participation. Specifically the research question addressed is to investigate "*to what extent has the decline in the earnings of low-skilled men contributed to the increase in disability payments*?". The disability programs examined are the Disability Insurance (DI) program and Supplemental Security Income (SSI). At the time of the writing of the paper these programs accounted $US51 billion in expenditure, so understanding what drives participation in these programs is key when building policy and spending forecasts. Overall the paper finds that DI responds to earnings with an elasticity of 0.3 to 0.4 and that SSI responds to earnings with an elasticity of 0.4 to 0.7. As hypothesised, this suggests that disability scheme participation is responds positively to earnings decreases.

Motivation

Previous work by Donald O. Parsons (1980) explored how DI payments contributed to a lower labour market participation rate. Parsons hypothesized that the highly progressive disability payments have causes the rise in labour force non-participation among low-wage men from 4.2% to 8.4% during 1942 to 1976, despite the payments only being available to those with physical disabilities. Later, Bound and Waidmann (1992) argued that more advanced healthcare has caused an increase in disabled survivors who require DI payments and cannot work, suggesting that the causal link suggested by Donald O. Parsons doesn't hold. In this paper Black, Daniel and Sanders use a different approach, looking at how long term changes in earnings affect participation in disability programs.

Program and Method Overview

The authors use the coal boom and bust (1970-1993) as a natural experiment to provide exogenous changes in long term wages, where the impact on wages was dependant on the size of the coal endowment of each area. From 1970 to 1977 the price of coal rose dramatically, then fell over a longer coal bust from 1983 to 1993.

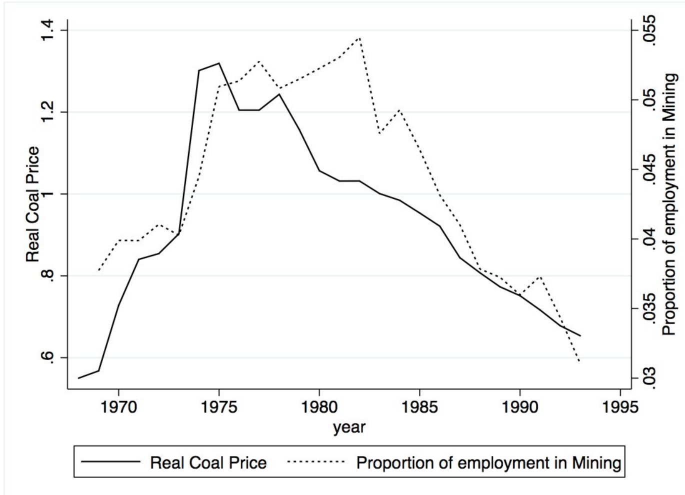

Figure 1: The real price of coal (1969-1993), mean proportion of employment in mining (1970-1993)

The authors use data from Kentucky, Ihio, Pennsylvania and West Virginia, states with large coal seams. They classify counties by the size of their coal reserves, making the inference that the magnitude of the effect of the coal price changes is proportional to the size of a county's coal reserves. A county is defined as a large coal county if it has at least a billion tons of coal reserves, a moderate coal county has between 100 million and a billion tons, and a no coal county has less than 100 million tons. During the boom period the wages of those in counties with large reserves of coal grew 3.5% faster than non-coal producing counties, while during the bust wages grew 3.4% faster in non-coal counties.[1] Thus the effects of the coal price changes were not distributed evenly across all counties, allowing the authors to use this as a natural experiment. This can be seen graphically in figure 1.2, which shows how wages and DI react to coal price changes in large coal and no coal areas. No coal areas show no correlation, while large coal areas show a positive correlation between earnings and coal price, and a negative correlation between earnings and DI payments, supporting the author's hypothesis.

These exogenous changes in earnings allowed the authors to measure the reactions of DI and SSI payments. In alignment with their research question, the coal boom and bust focuses the paper on low skilled men in coal production, a demographic that has struggled with to the depression of their wages due to globalization and changing commodities prices.

The study focuses on changes in long term earnings, as eligibility for the DI and the majority of the SSI program require the individual to be unable to work for at least 12 months due to a physical impairment and have been out of work for at least 5 months. These requirements make entering the disability programs a long term choice, one

[1] Taken from table 6 of the paper.

influenced by long term earning prospects rather than short term wage fluctuations or temporary unemployment. Changes in earnings induced by changes in the price of coal are likely due to the opening, expansion or closure of mines, and thus reflect long term changes in earnings, making the natural experiment suitable to induce earnings changes that are relevant to DI and SSI decisions.

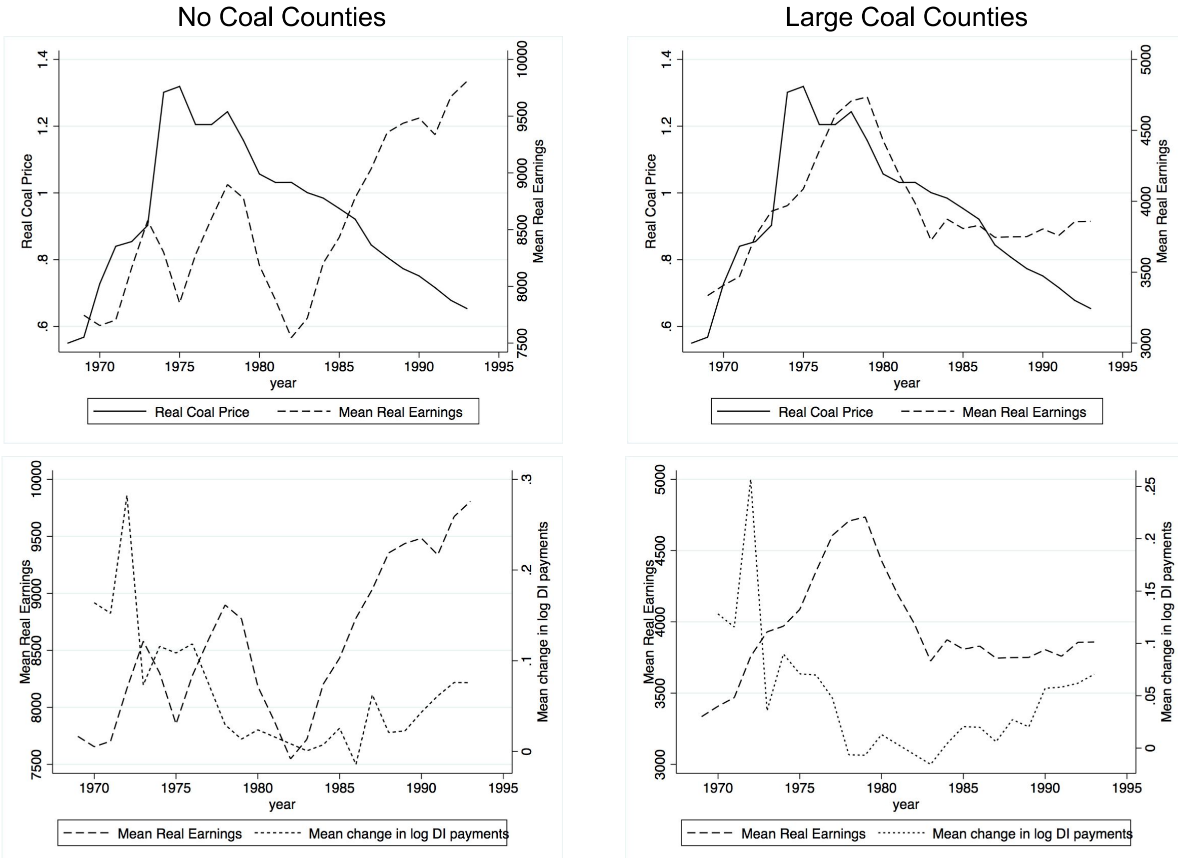

Figure 1.2: real coal price, real earnings and change in log DI payments by region, (1969-1993)

Data and Summary Statistics

The paper uses county level panel data for 1969 to 1993, as gathering individual level data was impractical. The data used covers Kentucky, Ohio, Pennsylvania and West Virginia. The use of county level data allows the authors to correlate intra-state variation in earnings and disability program payments while controlling for inter-state variations. All counties are not alike, with some having far larger populations than others. This is accounted for by the authors through adding in controls for population size, dense population centers[2] and through using the log difference for each wage and DI / SSI variable which normalizes the counties absolute size through:

$$Log(x_t) - Log(x_{t-1}) = Log(\frac{x_t}{x_{t-1}}) \approx \frac{x_t}{x_{t-1}} + 1 = \frac{1}{100}\ percentage\ change\ in\ x$$

[2] Accounted for by the msa dummy variable, which is true if the county contains a metropolitan statistical area - an area of high population density.

This approximation holds for small changes in x, so we can interpret the log differenced variables as their percentage change divided by 100, a rate of growth.

Table 2 - Summary Statistics by coal price period and coal reserves

Variables	All Counties	Large Coal Counties	Moderate Coal Counties	No Coal Counties
Coal Boom (1970-1977)				
Logarithmic difference in SSI payments	0.063	0.061	0.071	0.06
	(0.217)	(0.189)	(0.212)	(0.225)
Logarithmic difference in DI payments	0.128	0.102	0.119	0.136
	(0.087)	(0.08)	(0.08)	(0.09)
Logarithmic difference in county earnings	0.03	0.058	0.034	0.022
	(0.081)	(0.078)	(0.079)	(0.081)
Logarithmic difference in population	0.013	0.015	0.012	0.012
	(0.018)	(0.018)	(0.021)	(0.018)
Logarithmic difference in real price of coal	0.094	0.094	0.094	0.094
	(0.140)	(0.140)	(0.140)	(0.140)
Logarithmic difference in coal value instrument	0.252	0.721	0.547	0.049
	(0.648)	(1.080)	(0.823)	(0.223)
Million tons coal 20+ seam	457.855	2562.68	412.444	6.427
	(1107.692)	(1779.475)	(256.616)	(19.383)
Fraction of economy in manufacturing	0.272	0.172	0.284	0.289
(1969)	(0.159)	(0.151)	(0.167)	(0.149)
MSA indicator	0.261	0.191	0.296	0.264
	(0.439)	(0.394)	(0.457)	(0.441)
Population in thousands (000)	84.387	59.148	80.397	91.318
	(193.684)	(71.648)	(190.995)	(211.832)
Coal Bust (1983-1993)				
Logarithmic difference in SSI payments	0.062	0.067	0.063	0.061
	(0.073)	(0.07)	(0.068)	(0.075)
Logarithmic difference in DI payments	0.033	0.030	0.028	0.035
	(0.091)	(0.090)	(0.099)	(0.089)
Logarithmic difference in county earnings	0.017	-0.009	0.008	0.026
	(0.077)	(0.086)	(0.058)	(0.079)
Logarithmic difference in population	0.002	-0.007	-0.002	0.005
	(0.013)	(0.012)	(0.012)	(0.013)
Logarithmic difference in coal value instrument	-0.111	-0.318	-0.241	-0.022
	(0.149)	(0.137)	(0.107)	(0.057)
Logarithmic difference in real price of coal	-0.041	-0.041	-0.041	-0.041
	(0.018)	(0.018)	(0.018)	(0.018)
Population in thousands (000)	85.659	58.476	78.259	94.163
	(179.036)	(68.592)	(170.078)	(197.3)
Counties	330	47	71	212

Econometric Model: Semiparametric Models

The authors aim to estimate the causal effect of earnings on DI and SSI payments, B_1.

$$\Delta y_i = B_0 + B_1 \Delta earnings_i + \varepsilon_i$$

Where Δy_i is the log difference in DI or SSI payments. They use the log differenced variables, so the coefficient on earnings can be interpreted as the elasticity of DI and SSI payments with respect to earnings. There are interstate variations in how individuals qualify DI and SSI that could bias the estimates, and as such the authors only focus on intrastate variation in the variables of interest, using state level controls to remove interstate differences. After controlling for interstate variations and year fixed effects, the econometric specification takes the form:

$$\Delta y_{ist} = B_0 + B_1 \Delta earnings_{ist} + B_{2st} year_{st} + B_3 x_{ist} + \varepsilon_{ist} \quad (OLS)$$

Where $\Delta earnings_{ist}$ is the log difference in earnings for county *i* in state *s* in time *t*, $year_{st}$ is a dummy variable for state *s* in year *t*, and x_{ist} is a vector of control variables. Using differenced variables removes any time invariant county specific heterogeneity, eliminating the need to use a fixed effects model. Furthermore, the use of year-state dummy variables accounts for county invariant time heterogeneity as well as interstate differences.

This OLS model doesn't use the coal price change directly, instead it correlates the changes in earnings during the boom and bust period to changes in DI and SSI payments without any attention to the causes of the change in earnings. As a result, this specification doesn't differentiate between short term earnings fluctuations and the long term earnings changes the authors are primarily concerned with, likely causing bias towards zero due to excess noise in earnings. This method is a poor choice in estimating the effect of earnings on DI and SSI.

To extract the movements in earnings due to the coal price changes alone the authors employ a two-stage (2SLS) model. The first stage models changes in earnings in each county due to changes in the value of the coal reserves of the county. The change in value of coal reserves is also expressed in log form:

$$\Delta coalVal_{ist} = log(CoalReserves_{is})(log(RealPriceCoal_t) - log(RealPriceCoal_{(t-1)}))$$

Where $RealPriceCoal_t$ is the July CPI adjusted July coal price index for year *t*, and $CoalReserves_{is}$ is the amount of coal reserves in tons in county *i*, state *s*. This interaction of time invariant coal stocks with county invariant coal prices creates a measure of coal value that varies over both counties and time. Coal value and two lagged values are used as an instrument for earnings, to account for the time lags associated with expanding coal production. This gives the first stage regression:

$$\widehat{\Delta earnings}_{ist} = \beta_o + \beta_1 \Delta coalVal_{ist} + \beta_2 \Delta coalVal_{is(t-1)} + \beta_3 \Delta coalVal_{is(t-2)}$$
$$+ B_{4st} year_{st} + B_5 x_{ist} + \varepsilon_{ist} \quad (first\ stage)$$

These long-term, coal-value induced changes in earrings, $\widehat{\Delta earnings}_{ist}$ are then used to model changes in DI and SSI payments:

$$\Delta y_{ist} = B_0 + B_1 \widehat{\Delta earnings}_{ist} + B_{2st} year_{st} + B_3 x_{ist} + \varepsilon_{ist} \quad (2SLS)$$

If coal value changes only influence long term earnings, this specification removes the short term fluctuations that prove problematic in the OLS estimates, allowing the 2SLS model to provide a more accurate estimate for the coefficient on earnings. The authors estimate the OLS and 2SLS models with and without control variables, the results of which are replicated in table 3.

Table 3 – Impact of earnings changes on DI and SSI payments, OLS and 2SLS estimates.

	OLS	2SLS	OLS	2SLS
Control Variables	Yes	Yes	No	No
DI payments (N = 7,260)				
Change in county earnings	-0.002	-0.345	0.002	-0.347
	(0.111)	(3.943)	(0.098)	(4.238)
SSI payments (N = 7,904)				
Change in county earnings	-0.022	-0.710	-0.020	-0.636
	(1.559)	(5.310)	(1.403)	(5.466)

Note: All regressions include year-state fixed effects. Absolute t-statistics are reported in parenthesis. Control variables used (where yes) are if the county is in MSA, log population, difference in log population, fraction of earnings from manufacturing (1969). See Replication Comparison for deviation from original.

As expected, the OLS results suggest a weak link between earnings changes and DI and SSI, with elasticities of disabilities payments to earnings of between -0.002 and 0.002 for DI and -0.020 and -0.022 for SSI, all of which are insignificant at the 95% confidence level. The ranges are due to the inclusion of the control variables.

The 2SLS results suggest a large and statistically significant result for DI payments, with an estimated elasticity of -0.345 to -0.347. SSI shows even larger results, with estimated elasticities of -0.710 to -0.636, which are again statistically significant.

The authors explore factors that may cause these estimates to be unreliable. Coal producing counties are more likely to be rural, where populations are lower and the doctor-patient is likely to be more intimate. This connection may bias the doctors into being more lenient with their disability diagnosis when they know the economic prospects of the patient are unfavorable, leading to even larger increases in payments during times of low earnings growth and an overestimation of the elasticity. The authors test this by comparing the estimates between the predominantly rural larger coal counties and the more urban moderate coal counties, finding the two results similar for both DI and SSI. This result was confirmed for DI payments, but not for SSI as seen below. This suggests that rural communities may cause bias in SSI but not DI results.

Comparison of coefficient on earnings by regions, 2SLS model without controls

	Large and no-coal counties only	Moderate and no-coal counties only
DI payments	-0.344 (3.80) N=5,698	-0.263 (2.23) N=6,226
SSI payments	-0.625 (5.54) N=6,200	-0.680 (3.35) N=6,776

Note: Absolute t-statistics given in parenthesis. All regressions include year-state dummy variables. Standard errors were calculated using heteroskedasticity robust standard errors.

Another issue faces the SSI estimates. SSI is means tested, making workers ineligible if family income is too high, creating an incentive for non-disabled family members to avoid working to allow an eligible family member to collect SSI payments. As economic prospects improve, SSI payments may decrease due to both 1) eligible individuals seeking work and 2) non-eligible individuals seeking work, both of which contribute to the relationship between SSI and earnings. This is left untested as individual level data is unavailable.

Econometric Model: Nonparametric Models

The authors use a non-parametric 2SLS model to further ensure that only long term earnings growth is used to model SSI and DI payments. This model has the same second stage as before, except the control variables are removed as they had little effect on the results in table 3.

$$\Delta y_{ist} = B_0 + B_1 \widehat{\Delta earnings}_{ist} + B_{2st} year_{st} + \varepsilon_{ist} \quad (2SLS\ nonparam)$$

The first stage uses the variation in earnings between large coal, moderate coal and no coal areas over time to model earnings.

$$\widehat{\Delta earnings}_{ist} = \beta_o + \beta_1 C_{ist} + B_{2st} year_{st} + \varepsilon_{ist} \quad (first\ stage)$$

The usual year-state dummy variables are used, while C_{ist} is a new dummy variable vector that indicates if county i is in a large, medium or no coal area, and if the year is in the coal boom, peak or bust.[3] This specification causes the change in earnings for a given year to be constant for counties in the same coal-reserve-size region and state, greatly reducing county specific short-term fluctuations in earnings. The estimates of this model are given in table 7. The authors explore variations in earnings, DI, and SSI by time period and coal reserves in table 6, where they report combinations of β_1 from

[3] Thus C is a vector of 9 dummy variables.

the following regression, where Y_{ist} is earnings, DI, or SSI and C_{ist} is the same as above.

$$\Delta Y_{ist} = \beta_0 + \beta_1 C_{ist} + B_{2st} year_{st} + \varepsilon_{ist}$$

Table 6 - Estimates of the impact of the coal shocks on earnings, DI payments, and SSI payments

Period	Difference in Logarithmic difference in county earnings between:		Difference in Logarithmic difference in DI payments between:		Difference in Logarithmic difference in SSI payments between:	
	moderate and no coal reserve counties	large and no coal reserve counties	moderate and no coal reserve counties	large and no coal reserve counties	moderate and no coal reserve counties	large and no coal reserve counties
Boom	0.014	0.035	-0.016	-0.030	-0.004	-0.017
	(3.958)	(6.477)	(5.470)	(7.671)	(0.851)	(3.562)
Between	0.004	0.002	-0.010	-0.024	0.007	0.001
	(1.001)	(0.299)	(3.356)	(6.557)	(1.929)	(0.178)
Bust	-0.018	-0.034	-0.003	-0.004	0.004	0.014
	(6.584)	(8.195)	(1.603)	(1.307)	(1.687)	(4.281)

Note: Absolute t-statistics given in parenthesis. All regressions include year-state dummy variables. Standard errors were calculated using heteroskedasticity robust standard errors.

From table 6 it is clear that large and moderate coal county's earnings were affected more positively than no-coal counties during the boom period, with earnings growth that was 3.5% and 1.4% greater then no coal counties respectively, while during the bust earnings grew 3.4% and 1.8% slower. DI payments were 3% and 1.6% lower for large and moderate coal counties during the boom, while SSI payments where grew 1.7% and 0.4% slower during the boom and grew 1.4 and 0.4% faster during the bust when compared to no-coal counties. These results lend credibility to the non-parametric 2SLS estimates given in table 7, as they show statistically significant differences in earnings and disabilities payments between coal regions and time periods.

Table 7 - Wald estimates of the impact of coal price shocks on DI and SSI payments

Period	Counties with no coal compared to counties with large coal reserves	Counties with moderate coal reserves compared to counties with large coal reserves	Counties with no coal compared to counties with moderate coal reserves	All counties
DI payments				
Boom	-0.828	-0.852	-1.296	-0.887
	(4.562)	(2.449)	(3.078)	(5.133)
	N=2,072	N=944	N=2,264	N=2,640
Peak	-2.128	-5.574	-2.120	-2.075
	(1.133)	(0.218)	(0.886)	(1.300)
	N=777	N=354	N=849	N=990
Bust	0.112	0.053	0.198	0.124
	(1.261)	(0.284)	(1.603)	(1.621)
	N=2,849	N=1,298	N=3,113	N=3,630
All years	-0.333	-0.353	-0.234	-0.324
	(4.371)	(2.124)	(2.290)	(4.847)
	N=5,698	N=2,596	N=6,226	N=7,260
SSI payments				
Boom	-0.479	-0.704	-0.241	-0.460
	(2.936)	(1.779)	(0.656)	(3.051)
	N=2,056	N=944	N=2,248	N=2,624
Peak	-0.267	3.014	1.921	1.733
	(0.113)	(0.474)	(1.019)	(0.929)
	N=1,295	N=590	N=1,415	N=1,650
Bust	-0.382	-0.453	-0.257	-0.371
	(3.936)	(2.028)	(1.857)	(4.359)
	N=2,849	N=1,298	N=3,113	N=3,630
All years	-0.424	-0.532	-0.209	-0.398
	(5.465)	(2.638)	(1.527)	(5.595)
	N=6,200	N=2,832	N=6,776	N=7,904

Note: All regressions include year-state fixed effects. Absolute t-statistics are reported in parenthesis.

The results in table 7 give the elasticities of DI and SSI payments with respect to earnings in column 4, while columns 1 to 3 give elasticities for subgroups of coal regions. For SSI payments the authors obtain similar elasticities of -0.460 and -0.371 for the boom and bust periods respectively. However, for DI payments the authors obtain elasticities of -0.887 for the boom period, an unrealistically high estimate, and a positive elasticity of 0.124 for the bust period, which is inconsistent with theory and previous results.

The authors suggest that the volatile estimates for DI may be due to a negative correlation between DI growth and coal reserves, which causes an overestimation of the elasticity during the boom period and an underestimation during the bust. This tested by including coal-region fixed effects in the non-parametric model, as in table 8, which yields more reasonable estimates for DI elasticity ranging from -0.374 to -0.413 over regions and between -0.385 and -0.408 over time periods[4]. These results are much more consistent than those found in table 7 where regional fixed effects were excluded. The authors also re-estimate the semi-parametric 2SLS estimates for DI from table 3 to include coal-region fixed effects, changing the elasticity estimate from -0.347 to -0.275.

Table 8 – Estimates of elasticity of DI with respect to earnings, with fixed effects

	Coefficient on differenced log earnings	Moderate coal fixed effect	Large coal fixed effect
(1) Full sample (N = 7,260)	-0.386	-0.010	-0.017
	(5.422)	(5.200)	(-6.350)
Two way comparisons by coal region			
(2) No coal and moderate coal reserve counties	-0.408	-0.011	
(N = 6,226)	(3.567)	(5.240)	
(3) No coal and large coal reserve counties	-0.374		-0.017
(N = 5,698)	(4.712)		(6.047)
(4) Moderate coal and large coal reserve counties	-0.413	0.008	
(N = 2,596)	(2.361)	(2.914)	
Two way comparisons by period			
(5) Peak and Bust	-0.408	-0.010	-0.018
(N = 4,620)	(3.517)	(3.942)	(4.188)
(6) Boom and Bust	-0.385	-0.010	-0.017
(N = 6,270)	(5.251)	(4.926)	(6.053)
(7) Boom and Peak	-0.295	-0.011	-0.02
(N = 3,630)	(1.291)	(3.179)	(2.973)
Other 2SLS estimates			
(8) Change in coal value and two lags as instruments, fixed effects for coal region.	-0.275	-0.010	-0.017
(N = 7,260)	(3.655)	(5.264)	(6.772)
(9) Change in coal value and two lags as instruments, county fixed effects. (N = 7,260)	-0.271		
	(3.644)		
(10) Period-Region interactions as instruments, county fixed effects (N = 7,260)	-0.386		
	(4.947)		

Notes: All regressions include year-state fixed effects. Absolute t-statistics are reported in parenthesis. The dependant variable is log difference in real DI payments for all above regressions.

[4] Excluding the estimate from (7) boom and peak as it is statistically insignificant.

Challenges to causal interpretations

The authors confirm their findings and add strength to a causal interpretation of the estimated elasticities using two tests on the following two assumptions.

1. DI and SSI are only correlated with changes in long term earnings.
2. The 2SLS specification used captures changes in long term earnings.

Firstly they check that DI and SSI are only correlated with the long term earnings they specify. The residuals from the first stage regression of earnings are extracted then included in the semi-parametric 2SLS regression[5] from table 3:

$$\Delta y_{ist} = B_0 + B_1 \Delta \widehat{earnings}_{ist} + B_{2st} year_{st} + B_3 resid_i + \varepsilon_{ist} \quad (2SLS)$$

As in the below table, they find that the coefficients on the residuals are insignificant for both DI and SSI, suggesting that the variation in earnings pertinent to DI and SSI payments had been successfully extracted by the 2SLS specification.

Residual analysis of 2SLS specifications for DI and SSI

	DI	SSI
Coefficient on residual from first stage regression of earnings	-0.009	-0.008
	(0.55)	(0.54)

Note: Cluster robust standard errors were used. All regressions include year-state fixed effects. Absolute t-statistics are reported in parenthesis.

Secondly they look at how unemployment insurance (UI), a short term unemployment benefit, reacts to their long term earnings specification. The authors hypothesise that UI should be explained by transitory earnings shocks, so if their 2SLS model is correctly extracting only long term earnings, it should fail to model UI. This is confirmed by the results in table 9.

Table 9 – Impact of earnings growth on UI

	OLS	2SLS
Log Change in county earnings	-0.566	0.031
	(9.061)	(0.180)

Note: Cluster robust standard errors were used. All regressions include year-state fixed effects. 2SLS regression uses coal region and time period interactions as instruments. Absolute t-statistics are reported in parenthesis.

These two tests improve confidence that the authors have estimated SSI and DI elasticities using the desired long-term earnings. This also explains the larger difference between OLS and 2SLS estimates. If DI and SSI payments are driven solely by long term earnings changes, the short term fluctuations in earnings introduces noise in earnings in the OLS estimate, biasing the estimated elasticities towards zero, as seen in table 3.

[5] The no control variable version is used in this case

Conclusion

In the paper the authors explored how disability program payments are influenced by economic conditions, using the coal boom and bust as a natural experiment to provide exogenous changes in long-term earnings that varied by time and location. They find that disability payments do react to earnings, with DI payments having an elasticity of between -0.3 and -0.4 with respect to earnings, and SSI payments having an elasticity of between -0.4 and -0.7, as shown in their semi and non-parametric estimates.

These results have important policy implications as they suggest that disability programs are an alternative to work for at least a subset of the disabled population and their families. The authors stress that this should be interpreted as showing the disabled individuals commitment to the workforce as opposed to taking advantage of the social security system. It could be added that it may be a simple rational decision to work when the large economic rewards compensate for the pain and forsaken disability payments, and to collect DI or SSI otherwise.

Replication Comparison

Figure 1	The figures match for the years available in the data set.
Table 2	The summary statistics match exactly or within 0.005 for all except for *Fraction of economy in manufacturing.* This is likely due to the authors using a different measure for earnings.
Table 3	OLS - Exact match to the authors coefficients and t-statistic. 2SLS - Exact match for DI, inexact but within 0.003 for SSI.
Table 6	Exact match to the authors
Table 7	Exact match to the authors
Table 8	Exact match to the authors for all regressions except for (9), (10) where the t-statistics do not match due to stata being unable[6] to calculate usual robust standard errors with the large number of fixed effects.
Table 9	OLS - near match to the authors 2SLS - exact match to the authors

[6] Stata returned “Warning: variance matrix is nonsymmetric or highly singular” if clustered errors were specified. Hence, plain “robust” errors were used.

Investigation of the Effect of the Coal Boom and Bust on Retail Sub-sectors

Introduction

The economic prospects of primary and secondary sector workers are often portrayed as being instrumental in the welfare of a region, justifying incentives to keep such industries operational. This paper aims to explore the importance of the primary sector to the tertiary retail sub-sectors.

The coal boom and bust of 1970s and 1980s caused large fluctuations in the earnings and employment of those in the coal industry in coal producing regions, allowing the effects of this primary sector on the retail sector to be estimated.

Black, Daniel and Sander (2002) found that while earnings fluctuations did insignificantly affect disabilities payments, long term earnings were a significant variable in determining disabilities scheme payments. This was hypothesised to be due to disabilities scheme participation being a long term decision, hence long term earnings changes would drive decision making. A similar prediction is made for retail spending and therefore retail income. Retail purchases have an associated lifespan that I assume is taken into account in the purchase decision. A drink at a bar may only last a matter of minutes, while a car is expected to last many years. Hence I expect long term earnings to be important to all retail spending, but to have a larger impact on goods with a longer lifespan.

Literature review

Black, Mckinnish and Sanders (2005) examine the general spillover effects of the coal boom and bust on entire sectors such as construction, retail, manufacturing, as well as poverty, population growth and wages. For the retail sector as a whole, they find that the growth in employment, wages and gross earnings were significantly positively impacted by the coal boom, and significantly negatively impacted by the coal bust. During the peak of the coal price cycle, when coal prices were relatively constant, they find that the difference in retail growth between coal and non-coal regions is insignificant, supporting the hypothesis that coal price growth causally affected the retail sector.

Testing for spillover - Non-parametric Model

Firstly, I aim to test for spillover from the coal price cycle into retail sub-sectors. A natural experiment approach is used where the counties are divided into two groups - those with low or no coal reserves and those with large coal reserves, using the same definitions as before. Using the reasoning of Black (2005), this creates a spatial gap between the large coal (treatment) and no coal (control) counties, reducing cross border economic spillover effects. This adds some believability to the implied assumption that a wage increase in a county only affects businesses in that county. This assumption is also stronger for retail than other industries, as clients are more likely to be consumers and hence local, as opposed to other businesses that could potentially operate in other counties in business-to-business sectors such as manufacturing. For all the models in this study, only large and no-coal counties are considered.

The following model is used to estimate the average difference in growth of retail sub-sector indicators between large and no-coal counties for each time period of the coal price cycle.

$$\Delta lnY_{ist} = B_0 + \sum_{T=1}^{3} \delta_T L_{is} P_{tT} + \sum_{T=1}^{3} \phi_T P_T + B_{1st} year_{st} + u_{ist}$$

Y is a retail sub-sector indicator, L_{is} *indicates if county i is in a large coal region,* P_{tT} *indicates if t is coal price period T, and* $year_{st}$ is the same year-state control variable used by Black (2002). δ_T can be interpreted as the average difference in growth of indicators between large and no-coal counties in price period T.

Table E1 – Average difference in growth in retail sector employment, wages and earnings between large and no coal regions

Retail Indicator, Y	Boom	Peak	Bust
All Retail Employment	0.012	0.002	-0.012
	(2.94)	(0.44)	(5.00)
Average Implied Retail Wages	0.004	-0.002	-0.004
	(1.65)	(0.88)	(3.28)
Earnings in:			
All Retail	0.016	0.000	-0.016
	(4.865)	(0.020)	(6.344)
Groceries	0.006	-0.011	-0.013
	(1.209)	(1.919)	(3.263)
Home Furnishings	0.012	-0.002	-0.026
	(1.774)	(0.203)	(3.142)
Eating, Drinking	0.020	0.007	-0.008
	(3.040)	(0.845)	(2.045)
Building Materials	0.020	-0.016	-0.006
	(2.031)	(1.648)	(0.888)
Cars	0.010	0.005	-0.014
	(2.248)	(0.945)	(2.967)
Clothing	0.010	-0.003	-0.013
	(1.202)	(0.189)	(1.413)

Note: Absolute t-statistics given in parenthesis. All regressions include year-state dummy variables. Standard errors were calculated using heteroskedasticity robust standard errors.

As in Black (2005), the indicators show insignificant differences in growth during the peak period. The retail sector as a whole experienced significantly more growth in large coal counties during the boom and significantly less in large coal counties during the bust, as shown by growth in employment, earnings, and in the bust, implied wages as well.

Retail sub-sectors in large coal counties show the same pattern of more earnings growth then no-coal counties during the boom and less during the bust, with varying degrees of significance. The exception is clothing earnings, which isn't significant in any period. As a result the clothing sub-sector is largely excluded from further discussion. This natural experiment suggests that the earnings from the coal sector significantly affect most retail sub-sector earnings.

Estimating Spillover Magnitude - Parametric Model

Having shown that the retail sub sectors are affected by coal earnings, we now aim to estimate the magnitude of this affect. Due to the hypothesis that long term earnings are more useful than short term fluctuations in determining spending, two models are estimated.

The OLS model simply estimates the growth in retail sub-sector earnings, Y_{ist} as a function of growth in mining earnings, M_{ist}, using $year_{st}$ to control for time trends and inter-state differences. Further control variables are not included due to their insignificance to the results in Black (2002) and to the results in these models.

$$\Delta lnY_{ist} = B_0 + \delta_{OLS}\Delta lnM_{ist} + B_{1st}year_{st} + u_{ist} \quad (OLS)$$

To extract long term mining earnings growth $\Delta ln\widehat{M}_{ist}$ the same 2SLS approach as Black (2002) is used.

$$\Delta ln\widehat{M}_{ist} = \beta_o + \beta_1\Delta coalVal_{ist} + \beta_2\Delta coalVal_{is(t-1)} + \beta_3\Delta coalVal_{is(t-2)} + B_{4st}year_{st} + \varepsilon_{ist} \quad (first\ stage)$$

The long term earnings growth is then used to model growth in retail sub-sector earnings.

$$\Delta lnY_{ist} = B_0 + \delta_{2SLS}\Delta ln\widehat{M}_{ist} + B_{1st}year_{st} + u_{ist} \quad (2SLS)$$

$100\delta_{OLS}$ and $100\delta_{2SLS}$ can be interpreted as growth of retail sector earnings when mining earnings increase by 1%, hence δ_{OLS} and δ_{2SLS} are the elasticities of retail sector earnings with respect to mining earnings. Estimates are given in table E2.

Table E2 – Impact of growth in mining earnings on growth in retail earnings, OLS and 2SLS		
Retail Sector	OLS	2SLS
All Retail	0.005	0.134
	(2.152)	(3.394)
Groceries	0.005	0.116
	(1.338)	(1.824)
Home Furnishings	0.002	0.161
	(0.296)	(2.006)
Eating, Drinking	0.013	0.082
	(2.202)	(1.391)
Building Materials	0.007	0.190
	(1.069)	(1.461)
Cars	-0.004	0.137
	(0.653)	(2.580)
Clothing	0.014	0.032
	(1.558)	(0.262)

Note: Absolute t-statistics given in parenthesis. All regressions include year-state dummy variables. Standard errors were calculated using heteroskedasticity robust standard errors.

The first observation is that the elasticities are often small in magnitude yet significant. This is to be expected, as changing spending on retail is only one of many options a mining worker faces when evaluating how to allocate a change in earnings. Secondly, the 2SLS estimates are many times larger than the OLS estimates. I expect this is due to the additional noise in M_{ist} compared to $\widehat{M}_{ist}$ heavily biases the OLS estimates towards zero.

Considering the significant 2SLS elasticities, we see that they range from 0.134 to 0.161, while the elasticity of clothing is highly insignificant, the other retail sub sectors all demonstrate positive elasticities within the range 0.08-0.20 and with t-statistics over approximately 1.4. While this doesn't demonstrate strict statistical significance, it does show with a reasonable probability that retail earnings do react positively to mining earnings. For the retail sector as a whole, a highly significant elasticity of 0.134 is estimated. This shows some support for policy to support primary industries.

As $\Delta ln\widehat{M}_{ist}$ gives long term mining earnings growth, we can extract short term mining earnings growth $\Delta ln\ \widehat{m}_{ist}$ as a residual: $\Delta ln\ \widehat{m}_{ist} = \Delta ln\ M_{ist} - \Delta ln\ \widehat{M}_{ist}$

The elasticity of each retail sub-sector with respect to short and long term earnings can then be estimated. Results are reported in table E3.

$$\Delta lnY_{ist} = B_0 + \delta_{long}\Delta ln\widehat{M}_{ist} + \delta_{short}\Delta ln\ \widehat{m}_{ist} + B_{1st}year_{st} + u_{ist}\ \ (2SLS)$$

As $\widehat{m_{ist}}$ and $\widehat{M_{ist}}$ are orthogonal, this leaves this estimates on long term earnings unchanged.

Table E3 – Impact of growth in mining earnings on growth in retail earnings, separated long and short term earnings growth

Retail Sector	Long term earnings	Short term earnings
All Retail	0.134	0.004
	(3.394)	(1.715)
Groceries	0.116	0.005
	(1.824)	(1.113)
Home Furnishings	0.161	0.001
	(2.006)	(0.139)
Eating, Drinking	0.082	0.013
	(1.391)	(2.124)
Building Materials	0.190	0.005
	(1.461)	(0.857)
Cars	0.137	-0.005
	(2.580)	(0.835)
Clothing	0.032	0.014
	(0.262)	(1.550)

Note: 2SLS model used. Absolute t-statistics given in parenthesis. All regressions include year-state dummy variables. Standard errors were calculated using heteroskedasticity robust standard errors.

As expected, the magnitude of the elasticity of retail earnings with respect to long term earnings is many times greater than that with respect to short term earnings. This is consistent with mining workers making spending choices through evaluating their long term earnings prospects rather than short term fluctuations. Secondly, as hypothesised, goods with longer lifespans tend to exhibit a larger proportion of their overall elasticity with respect to mining earnings explained by long term earnings, as in figure E1.

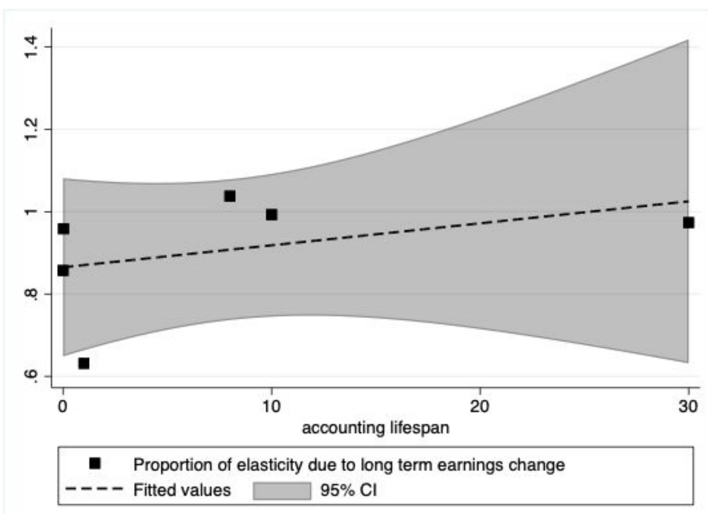

Figure E1 - Proportion of overall elasticity with respect to mining earnings due to long term earnings. Proportion = $\delta_{long}/(\delta_{long}+\delta_{short})$ to normalize between sub-sectors

This effect is hypothesised to exist for two reasons. Firstly, long term earnings changes allows for the justification of large purchases such as durable goods, while short term earnings changes are more irrelevant to such large savings outflows. This is supported as δ_{short} is small and insignificant for cars, home furnishings and building materials, retail goods with long lives. Secondly, long term and short term earnings increases allows for frivolous or luxury spending on goods with short lifespans. This is supported by the only significant elasticity for short term earnings being that of non-grocery food and drink.

Limitations

This analysis relies on the assumption that a mining wage increase in a county only affects businesses in that county, hence there is no significant spatial correlations between counties, in particular between regions of different coal endowment. While the focus on retail earnings makes this assumption more likely to hold, some spatial correlation is expected. This would cause an underestimation of the spillover effect and elasticities, as the models used only account for the effects on the county where the earnings change took place. As mining earnings are spent outside the county of work, the measured change in retail earnings is reduced for counties with larger earnings changes, while counties with low earnings changes can still experience increased retail earnings as money flows in from larger coal counties.

Conclusion

The coal boom and bust had a significant impact on the retail sector. The non-parametric estimations of the spillover effect of mining earnings into retail subsectors showed that during the boom retail earnings grew between 0.6% and 2% faster in areas with large coal reserves, and 0.6% and 2.6% slower during the coal bust.

Estimates of the elasticity of retail earnings with respect to mining earnings suggest that a 10% increase in mining earnings is associated with a 1.34% increase in retail earnings, an elasticity of 0.134, with sub-sectors showing elasticities between 0.082 and 0.190. Finally, its shown that long term earnings are the primary driver of spillover from the mining sector into retail earnings, although the product's lifespan plays a role, with longer lifespan products being relatively more responsive to long term earnings.

These results favor policy that sustains primary sector industries to maintain welfare. However, while significant, the magnitude of the effects of primary sector earnings on the retail sector was small, hence the costs of a policy could easily be greater than the benefits.

References

Black, Dan, Kermit Daniel, and Seth Sanders. "The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust." *American Economic Review* 92, no. 1 (2002): 27-50. doi:10.1257/000282802760015595.

Black, D., McKinnish, T., & Sanders, S. (2005). “The Economic Impact of the Coal Boom and Bust.” *The Economic Journal, no. 115* (2005), 449-476.

Bound, John, and Timothy Waidmann. "Disability Transfers and the Labor Force Attachment of Older Men: Evidence from the Historical Record." *The Quarterly Journal of Economics* 107, no. 4 (1992): 1393-419. doi:10.3386/w3437.

Parsons, Donald O. "The Decline in Male Labor Force Participation." *Journal of Political Economy* 88, no. 1 (1980): 117-34. doi:10.1086/260850.

Appendix - Code

```
*  ??????????????????????????
*  ??      Replication     ??
*  ?? Author: Grant Holtes ??
*  ??   Date: 20/09/2018   ??
*  ??????????????????????????

//Open and process data file to get all the required variables


*open data file
use "/Users/grantholtes/Documents/2018/AMM/2 Paper
      Replication/Data/bdsdata.dta", clear

*drop statewise sum rows
drop if sttot == 1

*define panel
*panel is strongly balanced
gen t = year - 1967
egen id = group(fips)
xtset id t

gen mining = bMining / bEmploy

*-------------------------------------------------------------------*
*-------------------------Variable Creation-------------------------*
*-------------------------------------------------------------------*
*------------------------------DUMMIES------------------------------*
*make boom, bust and between dummies
gen boom = 0
replace boom = 1 if year >= 1970 & year <=1977

*between boom & bust is 'btwbb'
gen btwbb = 0
replace btwbb = 1 if year >= 1978 & year <=1982

gen bust = 0
replace bust = 1 if year >= 1983 & year <=1993

gen all = 0
replace all = 1 if year >= 1970 & year <=1993

gen period = 1 if boom
replace period = 2 if btwbb
replace period = 3 if bust
```

```
*make coal reserve
gen noCoal = 0
replace noCoal = 1 if coalres < 100
label variable noCoal "No Coal County"

gen modCoal = 0
replace modCoal = 1 if coalres >= 100 & coalres < 1000
label variable modCoal "Moderate Coal County"

gen larCoal = 0
replace larCoal = 1 if coalres >= 1000
label variable larCoal "Large Coal County"

gen region = 0
replace region = 1 if modCoal
replace region = 2 if larCoal

*Fraction of economy in manufacturing
*desired = 0.267 for all counties
gen fracmanu = Manufact / EarnPOW if year == 1969
replace fracmanu = fracmanu[_n-1] if fracmanu ==.
label variable fracmanu "Fraction of economy in manufacturing (1969)"

*$-----------------------REAL DOLLAR VALUES-------------------------$

*Real SSI, DI
*deflate using pyr CPI indicies
gen rSSI = dSSI / p7
gen rDI = pay / p7
gen rUI = dUnemply / p7


*Real Earnings TODO EarnPow is best so far, then Wages
gen rearn = EarnPOW / p7

*real coal price
gen rpcoal = pcoal7/ p7
gen rpoil = poil7 / p7


*-------------make / convert log difference variables---------------------*
* pay = DI
* dSSI = SSI ??
* PrivPOW = earnings     TODO or EarnPriv
local dlvars "rSSI rDI rearn Pop rpcoal rUI"
foreach x of local dlvars {
```

```
      gen l`x' = log(`x')
      gen dl`x' = D.l`x'
//    drop l`x'
}

label variable dlrSSI "Logarithmic difference in SSI payments"
label variable dlrDI "Logarithmic difference in DI payments"
label variable dlrearn "Logarithmic difference in county earnings"
label variable dlPop "Logarithmic difference in population"
label variable dlrpcoal "Logarithmic difference in real price of coal"

*Coal real value instrument
*"we first must specify a measure for the value of a county’s coal reserves.
*We use the log differences in the real price as the measure of price changes.
*This measure of the change in coal prices is multiplied by the log of coal
      reserves."
gen dlrcoalVal = log(coalres) * dlrpcoal
replace dlrcoalVal = 0 if dlrcoalVal == .
label variable dlrcoalVal "Logarithmic difference in coal value instrument"

gen dlrcoalValL1 = l.dlrcoalVal
gen dlrcoalValL2 = l.dlrcoalValL1

*------------------------REORDER DATASET--------------------------------*
order id year t msa boom bust larCoal modCoal noCoal dlrSSI dlrDI dlrearn
      dlrpcoal dlrcoalVal dlrcoalValL1 dlrcoalValL2 Pop fracmanu
sort id t

*EXTENSION*
*Earnings
gen nonMine = EarnPOW - Mining

local retailEarn "nonMine Mining Retail RetFood HomeFurn EatDrink BldgMat
      RetAuto RetApprl"
foreach x of local retailEarn {
      *get real values
      gen r`x' = `x' / p7
      *construct log difference
      gen lr`x' = log(r`x')
      gen dlr`x' = D.lr`x'

}
*Employment
local Employment "bRetail bMining"
foreach x of local Employment {
      gen L`x' = log(`x')
      gen dl`x' = D.L`x'
}
```

```
*Implied average "wages" - may relect more hours, not wage.
*Earnings per person is more realsitic
gen rMineWage = Mining / (p7 * bMining)
gen rRetWage = Retail / (p7 * bRetail)
local wage "rMineWage rRetWage"
foreach x of local wage {
	gen l`x' = log(`x')
	gen dl`x' = D.l`x'
}

gen LarBoom = 0
replace LarBoom = 1 if larCoal & boom
gen LarPeak = 0
replace LarPeak = 1 if larCoal & btwbb
gen LarBust = 0
replace LarBust = 1 if larCoal & bust

gen dlrMiningWL1 = dlrMining * WL1

gen drMining = D.rMining
gen drnonMine = D.rnonMine


//Estout summary stats
//Grant Holtes

*Define variables to include
local fullvarlist dlrSSI dlrDI dlrearn dlPop dlrpcoal dlrcoalVal coalres
	fracmanu msa Pop
local subvarlist dlrSSI dlrDI dlrearn dlPop dlrcoalVal dlrpcoal Pop

*Boom
estpost sum `fullvarlist' if boom
est sto ss1

estpost sum `fullvarlist' if boom & larCoal
est sto ss2

estpost sum `fullvarlist' if boom & modCoal
est sto ss3

estpost sum `fullvarlist' if boom & noCoal
est sto ss4

esttab ss1 ss2 ss3 ss4 ///
using "/Users/grantholtes/Documents/2018/AMM/2 Paper
	Replication/Output/BoomSS.csv", ///
 cell(mean(fmt(%9.3f)) sd(par fmt(%9.3f))) collabels(none) label ///
```

```
 mtitle("All Counties" "Large Coal Counties" "Moderate Coal Counties"  "No Coal
      Counties" ) ///
 plain replace

*Bust
estpost sum `subvarlist' if bust
est sto ss5

estpost sum `subvarlist' if bust & larCoal
est sto ss6

estpost sum `subvarlist' if bust & modCoal
est sto ss7

estpost sum `subvarlist' if bust & noCoal
est sto ss8

esttab ss5 ss6 ss7 ss8 ///
 using "/Users/grantholtes/Documents/2018/AMM/2 Paper
      Replication/Output/BustSS.csv", ///
 cell(mean(fmt(%9.3f)) sd(par fmt(%9.3f))) collabels(none) label ///
 mtitle("All Counties" "Large Coal Counties" "Moderate Coal Counties"  "No Coal
      Counties" ) ///
 plain replace

//esttab ss1 ss2 ss3 ss4 ss5 ss6 ss7 ss8, cell(mean(fmt(%9.3f)) sd(par
      fmt(%9.3f))) collabels(none) label mtitle("Boom: All Counties" "Boom:
      Large Coal Counties" "Boom: Moderate Coal Counties"  "Boom: No Coal
      Counties" "Bust: All Counties" "Bust: Large Coal Counties" "Bust:
      Moderate Coal Counties"  "Bust: No Coal Counties")

//Semi-Paramatic Regressions
//Grant Holtes

//Table 3

global Controls_ols msa lPop dlPop fracmanu i.year#i.state dlrearn
global NoControls_ols i.year#i.state dlrearn
global Controls_2SLS msa lPop dlPop fracmanu i.year#i.state (dlrearn =
      dlrcoalVal dlrcoalValL1 dlrcoalValL2)
global NoControls_2SLS i.year#i.state (dlrearn = dlrcoalVal dlrcoalValL1
      dlrcoalValL2)

*----------------OLS-----------------
*DI without controls =  0.002, T good
* 0.002
reg dlrDI $NoControls_ols,  vce(cluster id)
display _b[dlrearn]
```

```
display _b[dlrearn] / _se[dlrearn]

*DI with controls = -0.002, T ok
*-0.002
reg dlrDI $Controls_ols,  vce(cluster id)
display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]

*SSI without controls - close! = -0.021, T ok
* -0.020
reg dlrSSI $NoControls_ols, vce(cluster id)
display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]

*SSI with controls - good match! = -.022, T ok
* -0.023
reg dlrSSI $Controls_ols, vce(cluster id)
display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]

*---------------2SLS---------------
*regress fitted values
*DI without controls = -.347 - GOOD T
*□ -0.347
ivregress 2sls dlrDI $NoControls_2SLS, small vce(cluster id)
display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]


*DI with controls = -.345 GOOD T
* -0.345
ivregress 2sls dlrDI $Controls_2SLS if all, small vce(cluster id)
display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]


*SSI without controls - close! = -.636
* □-0.639

display _b[dlrearn]
display _b[dlrearn] / _se[dlrearn]
// //FOR MORE TESTS, F-Stat for first state
//ivreg2 dlrSSI $NoControls_2SLS, small ffirst

*SSI with controls - close! =   -.710
* -0.713
ivregress 2sls dlrSSI $Controls_2SLS, small vce(cluster id)
display _b[dlrearn]
```

```
display _b[dlrearn] / _se[dlrearn]

*TESTING FOR RURAL INFLUENCE
ivregress 2sls dlrSSI $NoControls_2SLS if (noCoal | larCoal), small vce(cluster
      id)
ivregress 2sls dlrSSI $NoControls_2SLS if (noCoal | modCoal), small vce(cluster
      id)
ivregress 2sls dlrDI $NoControls_2SLS if (noCoal | larCoal), small vce(cluster
      id)
ivregress 2sls dlrDI $NoControls_2SLS if (noCoal | modCoal), small vce(cluster
      id)

//Non-Paramatic Regressions
//Grant Holtes
set matsize 800 //"You're Gonna Need A Bigger (matrix)"
//TABLE 6
//Regression method - as in paper
matrix R = J(6,6,.)

local periods "boom btwbb bust"

local row = 1
foreach period of local periods {
            *Test earnings
            xtreg dlrearn larCoal modCoal i.year##i.state if `period',
vce(cluster id)
            *Mod Coal vs no Coal
            matrix R[`row',1]  = _b[modCoal]
            matrix R[`row'+1,1]  = _b[modCoal] / _se[modCoal]
            *lar Coal vs no Coal
            matrix R[`row',2]  = _b[larCoal]
            matrix R[`row'+1,2]  = _b[larCoal] / _se[larCoal]

            *test DI
            xtreg dlrDI larCoal modCoal i.year##i.state if `period',
vce(cluster id)
            *Mod Coal vs no Coal
            matrix R[`row',3]  = _b[modCoal]
            matrix R[`row'+1,3]  = _b[modCoal] / _se[modCoal]
            *lar Coal vs no Coal
            matrix R[`row',4]  = _b[larCoal]
            matrix R[`row'+1,4]  = _b[larCoal] / _se[larCoal]

            *test SSI
            xtreg dlrSSI larCoal modCoal i.year##i.state if `period',
vce(cluster id)
            *Mod Coal vs no Coal
            matrix R[`row',5]  = _b[modCoal]
```

```
            matrix R[`row'+1,5]  = _b[modCoal] / _se[modCoal]
            *lar Coal vs no Coal
            matrix R[`row',6]  = _b[larCoal]
            matrix R[`row'+1,6]  = _b[larCoal] / _se[larCoal]

            local row = `row' + 2
}
matrix list R //Print results

//TABLE 7
//2SLS method

matrix R = J(16,4,.)

local periods "boom btwbb bust all"

local row = 1
local row2 = 9
foreach period of local periods {
            *DI
            ivregress 2sls dlrDI i.year##i.state (dlrearn = i.region##i.period)
if `period' & (noCoal | larCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row',1]  = `diff'
            matrix R[`row'+1,1]  = `SE'
            //---------------------------------------------------------
            ivregress 2sls dlrDI i.year##i.state (dlrearn = i.region##i.period)
if `period' & (larCoal | modCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row',2]  = `diff'
            matrix R[`row'+1,2]  = `SE'
            //---------------------------------------------------------

            ivregress 2sls dlrDI i.year##i.state (dlrearn = i.region##i.period)
if `period' & (noCoal | modCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row',3]  = `diff'
            matrix R[`row'+1,3]  = `SE'
            //---------------------------------------------------------

            ivregress 2sls dlrDI i.year##i.state (dlrearn = i.region##i.period)
if `period', small vce(cluster id)
```

```
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row',4]  = `diff'
            matrix R[`row'+1,4]  = `SE'
            //--------------------------------------------------------

            * SSI
            ivregress 2sls dlrSSI i.year##i.state (dlrearn =
i.region##i.period) if `period' & (noCoal | larCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row2',1]  = `diff'
            matrix R[`row2'+1,1]  = `SE'
            //--------------------------------------------------------
            ivregress 2sls dlrSSI i.year##i.state (dlrearn =
i.region##i.period) if `period' & (larCoal | modCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row2',2]  = `diff'
            matrix R[`row2'+1,2]  = `SE'
            //--------------------------------------------------------

            ivregress 2sls dlrSSI i.year##i.state (dlrearn =
i.region##i.period) if `period' & (noCoal | modCoal), small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row2',3]  = `diff'
            matrix R[`row2'+1,3]  = `SE'
            //--------------------------------------------------------

            ivregress 2sls dlrSSI i.year##i.state (dlrearn =
i.region##i.period) if `period', small vce(cluster id)
            local SE = _b[dlrearn]/_se[dlrearn]
            local diff = _b[dlrearn]

            matrix R[`row2',4]  = `diff'
            matrix R[`row2'+1,4]  = `SE'
            local row = `row' + 2
            local row2 = `row2' + 2
}

matrix list R //Print results

//Table 8 - DI
```

```
matrix R = J(20,3,.)

*(1) Full sample
//TRY NEW INTERACTIONS FOR TAB7
*(1)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all, small vce(cluster id)
matrix R[1,1] = _b[dlrearn]
matrix R[2,1] = _b[dlrearn] / _se[dlrearn]

matrix R[1,2] = _b[modCoal]
matrix R[2,2] = _b[modCoal] / _se[modCoal]

matrix R[1,3] = _b[larCoal]
matrix R[2,3] = _b[larCoal] / _se[larCoal]


*(2)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all & (noCoal | modCoal), small vce(cluster id)
matrix R[3,1] = _b[dlrearn]
matrix R[4,1] = _b[dlrearn] / _se[dlrearn]

matrix R[3,2] = _b[modCoal]
matrix R[4,2] = _b[modCoal] / _se[modCoal]

*(3)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all & (noCoal | larCoal), small vce(cluster id)
matrix R[5,1] = _b[dlrearn]
matrix R[6,1] = _b[dlrearn] / _se[dlrearn]

matrix R[5,3] = _b[larCoal]
matrix R[6,3] = _b[larCoal] / _se[larCoal]

*(4)
ivregress 2sls dlrDI i.year##i.state modCoal (dlrearn = i.region##i.period) if
all & (modCoal | larCoal), small vce(cluster id)
matrix R[7,1] = _b[dlrearn]
matrix R[8,1] = _b[dlrearn] / _se[dlrearn]

matrix R[7,2] = _b[modCoal]
matrix R[8,2] = _b[modCoal] / _se[modCoal]

*(5)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all & (btwbb | bust), small vce(cluster id)
matrix R[9,1] = _b[dlrearn]
```

```
matrix R[10,1] = _b[dlrearn] / _se[dlrearn]

matrix R[9,2] = _b[modCoal]
matrix R[10,2] = _b[modCoal] / _se[modCoal]

matrix R[9,3] = _b[larCoal]
matrix R[10,3] = _b[larCoal] / _se[larCoal]

*(6)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all & (boom | bust), small vce(cluster id)
matrix R[11,1] = _b[dlrearn]
matrix R[12,1] = _b[dlrearn] / _se[dlrearn]

matrix R[11,2] = _b[modCoal]
matrix R[12,2] = _b[modCoal] / _se[modCoal]

matrix R[11,3] = _b[larCoal]
matrix R[12,3] = _b[larCoal] / _se[larCoal]

*(7)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn =
i.region##i.period) if all & (btwbb | boom), small vce(cluster id)
matrix R[13,1] = _b[dlrearn]
matrix R[14,1] = _b[dlrearn] / _se[dlrearn]

matrix R[13,2] = _b[modCoal]
matrix R[14,2] = _b[modCoal] / _se[modCoal]

matrix R[13,3] = _b[larCoal]
matrix R[14,3] = _b[larCoal] / _se[larCoal]

*(8)
ivregress 2sls dlrDI i.year##i.state modCoal larCoal (dlrearn = dlrcoalVal
dlrcoalValL1 dlrcoalValL2) if all, small vce(cluster id)
matrix R[15,1] = _b[dlrearn]
matrix R[16,1] = _b[dlrearn] / _se[dlrearn]

matrix R[15,2] = _b[modCoal]
matrix R[16,2] = _b[modCoal] / _se[modCoal]

matrix R[15,3] = _b[larCoal]
matrix R[16,3] = _b[larCoal] / _se[larCoal]

*(9)
ivregress 2sls dlrDI i.year##i.state i.fips (dlrearn = dlrcoalVal dlrcoalValL1
dlrcoalValL2) if all, robust
matrix R[17,1] = _b[dlrearn]
```

```
matrix R[18,1] = _b[dlrearn] / _se[dlrearn]

*(10)
ivregress 2sls dlrDI i.year##i.state i.fips (dlrearn = i.region##i.period) if
all, robust
matrix R[19,1] = _b[dlrearn]
matrix R[20,1] = _b[dlrearn] / _se[dlrearn]

matrix list R //Print results

//Checking residual series.
drop dlrearn_hat resid
regress dlrearn i.year#i.state dlrcoalVal dlrcoalValL1 dlrcoalValL2,
vce(cluster id)
predict dlrearn_hat
gen resid = dlrearn - dlrearn_hat
*Check DI
regress dlrDI i.year#i.state dlrearn_hat resid if all,  vce(cluster id)
*Check SSI
regress dlrSSI i.year#i.state dlrearn_hat resid if all,  vce(cluster id)

//TABLE 9
*Checking UI
regress dlrUI i.year#i.state dlrearn, vce(cluster id)
ivregress 2sls dlrUI i.year#i.state (dlrearn = i.region##i.period), vce(cluster
id) small

//Line graph of mining employment and coal prices
cd "/Users/grantholtes/Documents/2018/AMM/2 Paper Replication/Dofiles"
*Format variables so they match Summary Statistics
run FormatVars.do

label var rpcoal "Real Coal Price"

collapse (mean) mining rpcoal, by(year)

label var mining "Mean Proportion of Employment in Mining"
label var rpcoal "Real Coal Price"

twoway (line rpcoal year, yaxis(1)) (line mining year, yaxis(2)),
xtitle("year")

//Line graph of coal prices, earnings and DI
*reformat data
run FormatVars.do
collapse (mean)rpcoal dlrearn rearn rSSI rDI dlrSSI dlrDI, by(region year)
label var rpcoal "Real Coal Price"
label var rearn "Mean Real Earnings"
```

```
label var dlrDI "Mean change in log DI payments"
twoway (line rpcoal year, yaxis(1)) (line rearn year, yaxis(2))  if region==2,
xtitle("year")
twoway (line rearn year, yaxis(1)) (line dlrDI year, yaxis(2)) if region==2,
xtitle("year")

twoway (line rpcoal year, yaxis(1)) (line rearn year, yaxis(2))  if region==0,
xtitle("year")
twoway (line rearn year, yaxis(1)) (line dlrDI year, yaxis(2)) if region==0,
xtitle("year")

//(1) Average diff in growth between coal and noncoal areas
//(modCoal excluded to reduce geographic spillover)

//All of retail
*Employment
reg dlbRetail i.year#i.state LarBoom LarPeak LarBust boom bust  if (noCoal |
larCoal), vce(cluster id)
*Wages
reg dlrRetWage i.year#i.state LarBoom LarPeak LarBust boom bust  if (noCoal |
larCoal), vce(cluster id)

//Subsets of retail
*earnings
matrix R = J(14,3,.)
local row = 1
local retailSubsets "dlrRetail dlrRetFood dlrHomeFurn dlrEatDrink dlrBldgMat
dlrRetAuto dlrRetApprl"
foreach x of local retailSubsets {
      quietly reg `x' i.year#i.state LarBoom LarPeak LarBust boom bust  if
(noCoal | larCoal), vce(cluster id)

      local BOOM = _b[LarBoom]
      local BUST = _b[LarBust]
      local PEAK = _b[LarPeak]
      local BOOMt = _b[LarBoom]/_se[LarBoom]
      local BUSTt = _b[LarBust]/_se[LarBust]
      local PEAKt= _b[LarPeak]/_se[LarPeak]
      matrix R[`row', 1] = `BOOM'
      matrix R[`row', 2] = `PEAK'
      matrix R[`row', 3] = `BUST'
      matrix R[`row'+1, 1] = `BOOMt'
      matrix R[`row'+1, 2] = `PEAKt'
      matrix R[`row'+1, 3] = `BUSTt'
      local row = `row' + 2
}
matrix list R //Print results
```

```
//(1)
*(1.1)
*OLS and 2SLS
matrix R = J(14,2,.)
local row = 1
local retailSubsets "dlrRetail dlrRetFood dlrHomeFurn dlrEatDrink dlrBldgMat
dlrRetAuto dlrRetApprl"
foreach x of local retailSubsets {
	quietly reg `x' i.year#i.state dlrMining if (noCoal | larCoal),
vce(cluster id)
	matrix R[`row',1] = _b[dlrMining]
	matrix R[`row'+1,1] = _b[dlrMining] / _se[dlrMining]

	quietly ivregress 2sls `x' i.year#i.state (dlrMining=dlrcoalVal
dlrcoalValL1 dlrcoalValL2) if (noCoal | larCoal), small vce(cluster id)
	matrix R[`row',2] = _b[dlrMining]
	matrix R[`row'+1,2] = _b[dlrMining] / _se[dlrMining]

	local row = `row' + 2
}
matrix list R //Print results

*(2.2)
*2SLS - short term vs long term
quietly reg dlrMining dlrcoalVal dlrcoalValL1 dlrcoalValL2 i.year#i.state if
(noCoal | larCoal)
predict dlrMining_hat if (noCoal | larCoal)
gen resid = dlrMining - dlrMining_hat if (noCoal | larCoal)

matrix R = J(14,2,.)
local row = 1
local retailSubsets "dlrRetail dlrRetFood dlrHomeFurn dlrEatDrink dlrBldgMat
dlrRetAuto dlrRetApprl"
foreach x of local retailSubsets {
	//GET 2SLS results - valid as resid is ortogonal to dlrMining_hat, so
this is a easy way to get correct S.E.
	quietly ivregress 2sls `x' i.year#i.state (dlrMining=dlrcoalVal
dlrcoalValL1 dlrcoalValL2) if (noCoal | larCoal), small vce(cluster id)
	matrix R[`row',1] = _b[dlrMining]
	matrix R[`row'+1,1] = _b[dlrMining] / _se[dlrMining]

	quietly reg `x' i.year#i.state dlrMining_hat resid if (noCoal | larCoal),
vce(cluster id)

	matrix R[`row',2] = _b[resid]
	matrix R[`row'+1,2] = _b[resid] / _se[resid]

	local row = `row' + 2
```

```
}
matrix list R //Print results
```