Spring 2017

Essays in labor economics

Emily Catherine Leslie
University of Iowa

ESSAYS IN LABOR ECONOMICS

by

Emily Catherine Leslie

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Economics
in the Graduate College of
The University of Iowa

May 2017

Thesis Supervisor: Professor David Frisvold

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Emily Catherine Leslie

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Economics at the May 2017 graduation.

Thesis Committee: ______________________
David Frisvold, Thesis Supervisor

____
Padmaja Ayyagari

____
Alexandre Poirier

____
Joseph Price

____
Nicolas Ziebarth

ACKNOWLEDGEMENTS

I would like to thank my advisor, David Frisvold, for his support, genuine concern, valuable feedback, and encouragement throughout the research phase of my doctoral program. I am also grateful to the Economics Department at the University of Iowa for its financial support and for giving me the flexibility I needed as my family grew. Thank you to my friends at the BYU Economics Department, especially Lars Lefgren and Joseph Price, for their generosity in supporting me and cheering me on from my time as an undergraduate.

I am deeply grateful to my family. To my mother, Catherine Andersen, for showing me that a good education is valuable in every pursuit, and for helping care for my three precious children so that I could continue my studies. To my father, Eric Andersen, for his unwavering belief that I can do hard things, including finishing my degree. And especially to my husband, Russell Leslie, for supporting me without reservation in both word and action in more ways than I can possibly list.

ABSTRACT

This thesis considers how potentially vulnerable populations are affected by various economic and policy shocks. In the first chapter, I investigate the impact of natural resource booms on crime by estimating the effect of the coal boom and bust of the 1970s and 1970s on reported crime rates. I begin by demonstrating that changes in the value of coal reserves affected local economic conditions and population composition, both of which have theoretical and empirical links to crime. The net effect is theoretically ambiguous. The estimates suggest that the immediate impact of increasing the value of natural resources is to depress crime rates, primarily through changes in property crime, but these changes erode over time. My findings are consistent with an initial change in criminal activity in response to local labor market conditions that is subsequently offset by selective migration.

Individuals who are charged with committing a crime often find themselves behind bars while their case is adjudicated. In the United States, over 400,000 individuals are in jail each day waiting for their criminal cases to be resolved. The majority of these individuals are detained pretrial due to the inability to post low levels of bail (less than $3,000). In chapter 2, my coauthor and I estimate the impact of being detained pretrial on the likelihood of an individual being convicted or pleading guilty, and their sentence length, using data on nearly a million misdemeanor and felony cases in New York City from 2009 to 2013. Causal effects are identified using variation across arraignment judges in their propensities to detain defendants. We find that being detained increases the probability of conviction by over seven percentage points by causing individuals to plead guilty more often. Because pretrial detention is driven by failure to post bail, these adverse effects disproportionately hurt low-income individuals.

While some public policies create burdens that fall most heavily on low-income people and households, the public safety net is comprised of programs intended to protect and support this vulnerable population. In chapter three, my coauthors and I examine whether programs that provide vouchers to households can continue to influence behavior even after the household leaves the program. Using detailed scanner data, we test whether benefit vouchers received through the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) change household purchasing decisions and whether these changes continue to persist even after households are no longer eligible to participate in the program. In 2009, the package of goods available through WIC vouchers changed to include additional items and place nutritional restrictions on other items. Examining variation due to this package change, we show that the WIC vouchers change purchasing decisions consistent with the nutritional guidelines of the program. However, we find evidence of limited persistence post-eligibility, and that households exposed longer to the revised package are generally not more likely to continue to purchase these items after eligibility ends.

PUBLIC ABSTRACT

This thesis considers how potentially vulnerable populations are affected by various economic and policy shocks. In the first chapter, I investigate the impact of natural resource booms on crime by estimating the effect of the coal boom and bust of the 1970s and 1970s on reported crime rates. I begin by demonstrating that the boom and bust affected local economic conditions and population composition in counties with higher levels of coal reserves. Both labor market conditions and population composition have links to crime. My analysis suggests that the immediate impact of increasing the value of natural resources is to depress crime rates, primarily through changes in property crime, but these changes erode over time. My findings are consistent with an initial change in criminal activity in response to local labor market conditions that is subsequently offset by selective migration.

Individuals who are charged with committing a crime often find themselves behind bars while their case is adjudicated. In the United States, over 400,000 individuals are in jail each day waiting for their criminal cases to be resolved. The majority of these individuals are detained pretrial due to the inability to post low levels of bail (less than $3,000). In chapter 2, my coauthor and I estimate the impact of being detained pretrial on the likelihood of an individual being convicted or pleading guilty, and their sentence length, using data on nearly a million misdemeanor and felony cases in New York City from 2009 to 2013. We find that being detained increases the probability of conviction by over seven percentage points by causing individuals to plead guilty more often. Because pretrial detention is driven by failure to post bail, these adverse effects disproportionately hurt low-income individuals.

While some public policies create burdens that fall most heavily on low-income people and households, the public safety net is comprised of programs intended to protect and support this vulnerable population. In chapter three, my coauthors and I examine whether programs that provide vouchers to households can continue to influence behavior even after the household leaves the program. We test whether benefit vouchers received through the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) change household purchasing decisions and whether these changes continue after households are no longer eligible to participate in the program. We show that the WIC vouchers change purchasing decisions while households are eligible for the program. We find only short-lived persistence post-eligibility.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
RESOURCE BOOMS AND CRIME: EVIDENCE FROM THE MINING BOOM AND BUST OF THE LATE 1900s

1.1 Introduction

Crime in the United States generates societal costs. Anderson (1999) estimates the annual cost of crime to be about $1,100 billion. Widely adopted economic models such as Becker (1968) predict that better job market opportunities lower the relative benefit of crime, reducing criminal activity. On the other hand, if crime involves normal goods, opportunities for and rewards from criminal behavior may increase during economic expansions. Isolating the causal impact of economic conditions on crime is challenging. Reverse causality can arise if higher crime levels displace economic activity, or if criminal records reduce the job options of offenders.

In general, the empirical literature supports Becker's prediction that better labor market conditions result in lower crime rates, particularly for property crime. Kelly (2000), Raphael and Winter-Ebmer (2001), Gould et al. (2002), Machin and Meghir (2004), and Lin (2008) all find a positive relationship between unemployment and lower wages and crime. Several of these authors make use of instrumental variables to identify a causal effect (see Raphael and Winter-Ebmer (2001), Gould et al. (2002), and Lin (2008)). The existing literature also tends to exploit transitory changes in economic conditions. However, longer-term shifts in the labor market could plausibly have a different effect on criminal behavior, as affected individuals are likely to have different long-run expectations of their economic prospects. In addition, extrapolating from individual-level studies to make predictions about community-level crime rates will fail to take into account population composition effects driven by shifts in local economic conditions.

Resource booms are a special case of particular interest. They generate increased labor demand within specific industries that can influence wages and migration patterns, with the potential for spillovers to other industries. They are simultaneously transitory and unpredictable in duration. Growth in wages and job opportunities may reduce crime rates. At the same time, if resource booms induce selective migration (for example, attracting a large number of young males with low levels of education), affected communities may reasonably be concerned about whether the incoming groups are more prone to engaging in criminal behaviors than the average resident. The evidence on the impact of immigration on crime is mixed. The effects appear to be highly context-specific, with inter-country movement dominating the existing literature (see Bell et al. (2013),Chalfin (2014), Nunziata (2015)). The net impact of the labor market effects and the population composition changes on community crime rates can only be determined empirically.

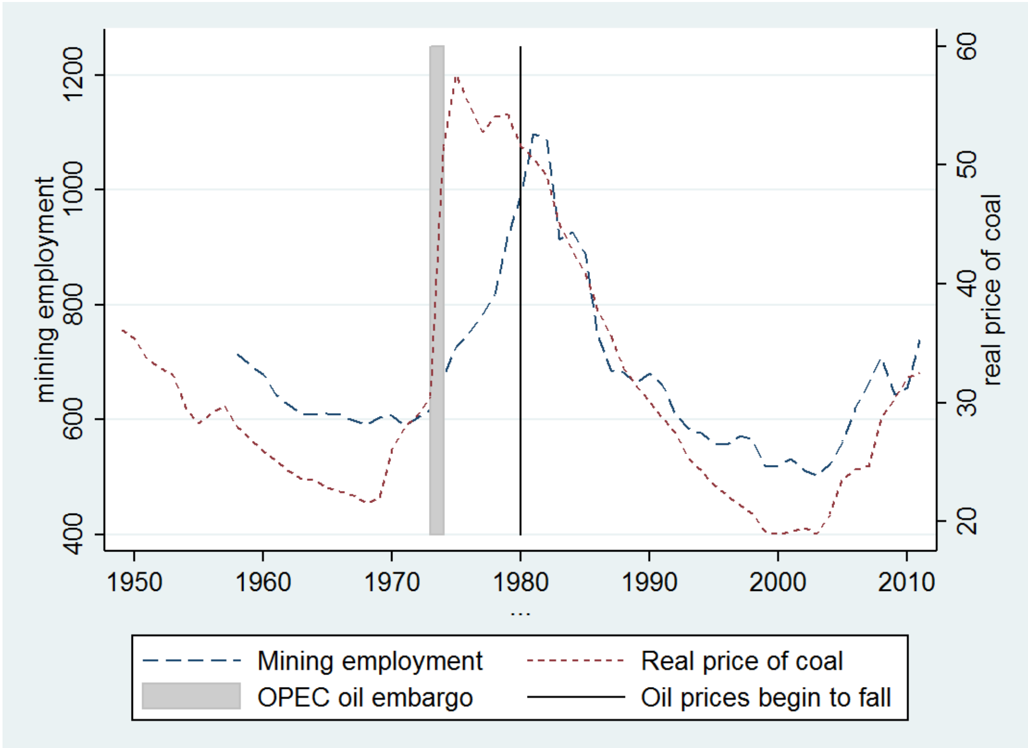
The mining boom and bust that took place in the United States during the 1970s and 1980s provides an ideal setting for investigating the impact of resource booms on crime. Global changes in the price of oil and the development of new technologies precipitated a marked increase in the price of coal followed by an equally noticeable decline. The entire boom/bust cycle took place within two decades. I use intertemporal variation in real coal prices combined with geographic variation in the distribution of coal to identify the impact of the resource boom/bust cycle on local labor market conditions, population composition, and crime rates. I find that an increase in the value of coal reserves increased earnings and jobs growth. Rising coal prices are also associated with larger populations and growing shares of young men. These effects are stronger during the bust than the boom. Young, unskilled men are particularly at risk for engaging in criminal activity (see Gould et al. (2002) and Freeman (2000)), but this relationship was more than offset by the impact of economic conditions. Increasing the value of coal reserves appears to reduce crime rates initially, but these changes erode over time.

1.2 Background

The OPEC oil embargo of the early 1970s drove up the price of oil, and the price of coal along with it, creating a dramatic increase in demand for workers in the mining industry. In the early 1980s, declining oil prices, new technologies reducing the number of workers needed to extract coal, and an increase in global coal production in response to the boom combined to bring down the price of coal (James, 1984). The result was a dramatic, decade-long increase in the number of workers employed in the mining industry followed by an equally dramatic collapse in mining employment. Figure 1.1 shows the rise and fall in the price of coal and the corresponding fluctuation in mining jobs. The powerful relationship between the coal and oil markets is apparent in the graph. The oil embargo coincides exactly with the most dramatic increase in the price of coal, and the year in which oil prices began to drop (1980) marks the beginning of a 20-year period of falling coal prices.

The empirical evidence reveals that some countries benefit from the presence of natural resources, while others, perversely, suffer as a result of their access to valuable commodities (Van der Ploeg, 2011). The literature is dominated by studies of countries with rich natural resources and weak institutions (see Angrist and Kugler (2008), Arezki and Van der Ploeg (2011), and Auty (2001)). The impact of spatial variation in the distribution to natural resource reserves within the United

Figure 1.1: The Real Price of Coal and National Mining Employment

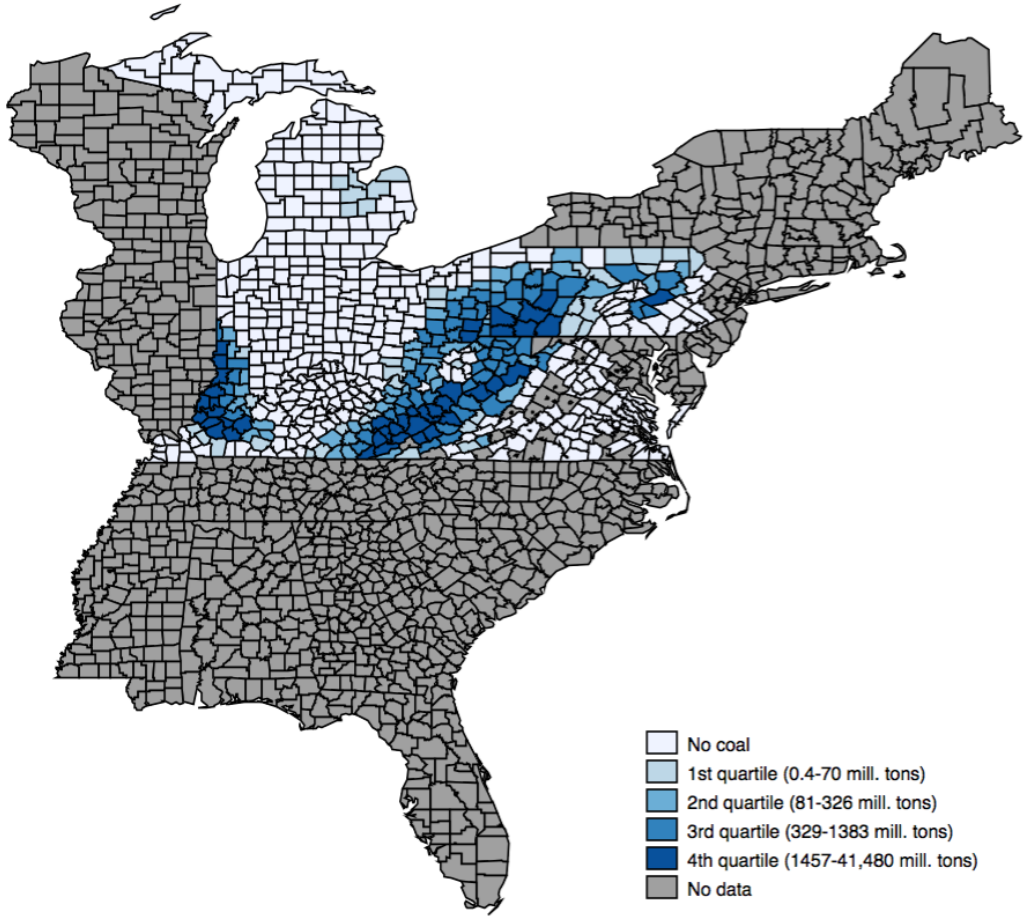


Notes: Mining employment is total mining employment in the country in 1000s of persons. The real price of coal is in 2005 dollars per short ton. National mining employment data come from the Current Employment Statistics (Establishment Survey), and data on the real price of coal come from the Energy Information Administration.

States has not been fully explored yet. Studies of the ongoing fracking boom in the United States suggest that wages and employment within the shale gas industry increased significantly, inducing more young men to drop out of high school (Cascio and Narayan, 2015). The evidence is mixed on whether these expansions in the shale industry were offset by reductions in other sectors (Cosgrove et al., 2015) or magnified by spillovers (Feyrer et al., 2015).

A series of papers by Black and coauthors examine the impact of the mining boom and bust on four Appalachian states (Kentucky, Ohio, Pennsylvania, and West Virginia). They estimate the effect of changes in the value of coal reserves on a variety of outcomes. Their results suggest that the mining boom increased mining employment, as well as creating local employment spillovers (Black et al., 2002)), decreased disability program participation, and decreased welfare expenditures, while the bust reversed these effects and led to outmigration, especially of young males (see Black et al. (2002), Black et al. (2003), and Black et al. (2005)).

Figure 1.2: The Distribution of Coal Reserves



Notes: Quartiles of coal reserves in eastern states for which data are available.

I build off of the methodology used in the existing literature to provide the first estimates of the impact of the coal boom on crime, and the first estimates employing a plausible identification strategy of the impact of natural resource-driven expansions on crime in the United States [1]. My

[1]Komarek (2014) compares crime trends in counties in the Marcellus region of New York and Pennsylvania with high levels of fracking activity with those that have low levels of fracking activity. His findings indicate that violent crime increased in fracking areas, but property crime did not.

sample includes the four states in the Black et al. sample as well as three additional states (Indiana, Michigan, and Virginia). Figure 1.2 displays the distribution of known coal reserves for the eight states east of the Mississippi River for which data are available, as described below. Coal counties in this area are spread throughout the eight states, and all, unsurprisingly, are contiguous with other coal counties. The map shows the largest concentrations of coal to be in the heart of Appalachia and in part of the Illinois basin, with a small pocket of counties with low levels of coal reserves in eastern Michigan. Most coal counties have fewer than half a billion short tons of known coal, but the right tail of the distribution is long, with the top quartile ranging from about one and a half billion to over 40 billion tons of coal.

1.3 Methodology

To create a picture of the impact of the mining boom and bust, I use a series of reduced form specifications to estimate how changing the value of coal reserves affected local economic conditions and crime rates. I prefer this approach to using the value of coal reserves as an instrumental variable because of the multiplicity of channels through which the rise and fall in coal prices could plausibly have affected crime rates. I could instrument for earnings, for example, as a proxy for the overall strength of the local economy, but it would be questionable to interpret the resulting coefficients as the impact of earnings on crime rates when the availability of jobs and population dynamics are also channels through which resource booms might affect criminal activity. Instead, I directly estimate the relationship between the degree to which a county was affected by the mining boom and bust and earnings, jobs, crime, and population characteristics. I follow Black et al. (2002) in constructing a measure of the change in the value of coal reserves:

$$\Delta v_{ct} = \ln(\text{coal}_c) * [\ln(p_t) - \ln(p_{t-1})] \tag{1.1}$$

where Δv_{ct} is the change in the value of coal reserves in county c from year $t-1$ to year t; coal_c is the amount of known coal in county c, measured in short tons; and p_t is the price of coal in year t, measured in 2005 dollars. For counties with no known coal reserves, I set $\ln(\text{coal}_c)$ equal to zero, as if they each had a single ton of coal[2]. Using the natural log of coal reserves is a logical choice given the skewness in the distribution of coal. Using the raw amount of known coal would result in estimates driven entirely by the extreme outliers in the long right tail. The regressions take the

[2]The average coal county in the sample has over 900 million short tons of coal.

following form:

$$\Delta y_{ct} = \beta_0 + \beta_1 \Delta v_{ct} + \beta_2 \Delta v_{ct-1} + \beta_3 \Delta v_{ct-2} + \beta_4 X_{ct} + \alpha_{st} + \epsilon_{ct} \tag{1.2}$$

where Δy_{ct} measures the change in an outcome of interest from year $t-1$ to year t. I estimate effects for the log change in total county earnings, the log change in the total number of jobs in a county, and the change in per capita reported crime rates[3]. The regressors of interest are the measure of the change in the value of coal reserves and two lags of this variable[4]. The lags are included to allow time for the industry response to rapid changes in coal prices, e.g., time to set up new mines. Figure 1.1 illustrates the time delay between the explosive growth in coal prices during the oil embargo and the expansion of the mining labor force. The vector of controls X_{ct} includes the log of county population and the log change in population from year $t-1$ to year t, as well as several county-level characteristics from the baseline year (1970): population per square mile, unemployment, share of the population below the federal poverty line, and net migration from 1960 to 1970. I control for the change in population, as well as pre-trends in migration, because changes in population may have a mechanical relationship with earnings and jobs. Also included are a vector of state-by-year fixed effects, α_{st}. Standard errors are clustered at the county level to correct for possible nonindependence.

To estimate population effects, I am restricted by the years in which population size and characteristics were measured. Fortunately, decennial census years match up well with the pre-boom (1970), peak (1980), and post-bust (1990) periods. I estimate two sets of regressions of the following form, one to measure effects during the boom, and one to measure effects during the bust:

$$\Delta y_c = \beta_0 + \beta_1 \ln \text{coal}_c + \beta_2 X_c + \alpha_s + \epsilon_c \tag{1.3}$$

---

[3]In Appendix A, I reestimate effects on earnings, jobs, and crime using an alternative treatment variable to relax the assumption implicit in the specification above that the relationships between the value of coal reserves and the outcomes of interest are linear. Instead of measuring the value of coal reserves, I use an indicator variable that takes a value of one for counties with coal reserves for years between 1974 and 1981, when the real price of coal was greater than \$50 dollars short ton, in 2005 dollars. This approach compares conditions in coal counties in peak price years to coal counties in off-peak price years and non-coal counties. The results are qualitatively similar.

[4]A model allowing for non-linearities in the level of coal, e.g., estimating effects by quartile, is one alternative to my specification. This approach would either require a larger number of regressors of interest, or would fail to utilize much of the variation in known coal reserves. My preferred method is parsimonious without ignoring potentially informative variation.

Δy_c is the change in population size or the share of a demographic group during the boom or bust decade. I estimate effects for the share of the population made up of young males (aged 18 to 30 years old), young females, older males (aged 31 to 50), and older females. The regressor of interest is the log of coal reserves, rather than the log of coal reserves interacted with the log change in price. Since the log change in price is the same for all counties during each decade, the price information would not capture useful variation. All specifications include controls for a set of county-level characteristics in 1970 (log of population, population per square mile, unemployment, and share of people below the poverty line), as well as state fixed effects. Robust (Huber-White) standard errors account for possible heteroskedasticity.

The identifying assumption in both equations (2) and (3) is that the distribution of coal is uncorrelated with factors other than the total value of coal reserves that influence changes in earnings, jobs, and crime rates, after conditioning on baseline measures of economic vitality and urbanness. It is reasonable to expect the presence of natural resources and the corresponding geological features to be correlated with the pre-boom levels of the outcomes of interest. I attribute the differential trends across counties with different levels of coal reserves entirely to the effect of the fluctuations in the value of coal. Additionally, the spike in coal prices increased incentives to find new coal reserves, so that the boom and bust could have directly affected exploration for and discovery of coal. By using the amount of known coal reserves pre-boom, I avoid endogeneity problems arising from the discovery of coal as a result of increased boom-driven economic activity.

1.4 Data

I bring together coal reserves data from several sources on counties from seven states east of the Mississippi (Indiana, Kentucky, Michigan, Ohio, Pennsylvania, Virginia, and West Virginia). Data on coal reserves for Kentucky, Ohio, Pennsylvania, and West Virginia were compiled for Black et al. (2002) (BDS) and made available by the authors. Coal reserves data for the other three states come from a series of reports published by the United States Geological Survey (USGS) between 1949 and 1953 on state coal resources (Cohee et al., 1950; Brown et al., 1952; Spencer, 1953). The BDS data are for coal from seams above a certain thickness. To make the USGS data comparable with the BDS data, I multiplied total coal reserves numbers by two-thirds. This adjustment is how Black, Daniel, and Sanders adjusted coal reserves measurements for states in their sample for which they did not have their preferred seam thickness breakdown (see their Data Appendix). There are 188 counties with known coal reserves according to these sources; 126 of the coal counties are in the Black et al. sample, and 28 come from the three additional states for which I have data[5]. The remaining 422 counties in these states for which I have earnings data are treated as having no known coal reserves.

Bureau of Economic Statistics Local Area Personal Income data provide information on annual county-level earnings for 1969 to 1993, which I convert to 2005 dollars. Baseline controls (population density, unemployment, fraction of residents below the federal poverty line, and net migration in the previous decade, all measured in 1970) come from the 1972 County Data Book. Twenty-three counties with no known coal reserves from Virginia for which I have earnings data do not appear in the 1972 County Data Book. This leaves a sample of 587 counties with sufficient data to estimate the impact of the mining boom and bust on earnings and population characteristics. Population from the decennial censuses of 1970, 1980, and 1990 are used to estimate population effects. County/year populations by sex and single year-of-age were compiled by the National Cancer Institute for the Surveillance, Epidemiology, and End Results Program (SEER) population database.

Data on crimes reported (all crimes, property crimes, and violent crimes) are from the FBI's Uniform Crime Reporting Program (UCR). UCR reporting is far from complete during the sample period, with more law enforcement agencies submitting reports over time. Analysis at the agency level is problematic. Many agencies have overlapping jurisdictions (e.g. one agency may cover a local university, and another the town in which the university is located). Some agencies are covered by other agencies in certain years. The report from the "covering agency" in that year adds the crime counts of the agencies for which it is reporting to its own crime numbers, making it impossible to generate separate statistics for each agency. Aggregated county-level crime rates based on UCR data have been used by other researchers. However, the existing aggregated datasets rely on inconsistent imputation methods. Especially problematic is the combination of crime data from the UCR with population data from the Census, which implicitly assumes that non-reporting agencies (a large set that changes over time) do not have any crime (Maltz and Targonski, 2002). I therefore create my own county-level dataset. The raw UCR data on monthly reported crimes, along with population data by agency, were generously provided by Michael Maltz. I convert this to yearly data, which requires some imputation for agencies that report aggregated data several months a year. For example, for states where the missing code for January through November indicates aggregation in December, I assume that the number of crimes reported in December represents the total for the

[5]Earnings data are available for all counties with known coal reserves

entire year. I use the population data included in the UCR data to calculate per capita crime rates, thus avoiding the problem of including populations for non-reporting agencies in the denominator. For the panel analysis, I restrict attention to agencies that reported (either themselves or through

Table 1.1: Summary Statistics

	Full sample	No coal	Coal
Coal reserves (mill. tons)	306.47	0.00	925.39
	(929.56)	(0.00)	(1429.68)
Baseline (1970)			
Net migration (1960-1970)	6.16	9.74	-1.07
	(12.74)	(12.54)	(9.75)
Pop. per square mile	224.41	270.16	132.03
	(856.29)	(1033.03)	(220.41)
Below FPL, 1970	14.65	12.32	19.37
	(10.31)	(7.25)	(13.51)
Boom (1970-1980)			
Log diff. total earnings	0.16	0.13	0.23
	(0.22)	(0.20)	(0.25)
Log diff. total jobs	0.17	0.16	0.19
	(0.16)	(0.16)	(0.18)
Log diff. population	0.12	0.12	0.11
	(0.11)	(0.12)	(0.09)
Change frac. young male	0.03	0.02	0.03
	(0.02)	(0.02)	(0.02)
Change in crimes reported (per 1000 people)	12.55	13.18	11.34
	(16.41)	(17.64)	(13.70)
Bust (1980-1990)			
Log diff. total earnings	0.15	0.21	0.01
	(0.23)	(0.20)	(0.21)
Log diff. total jobs	0.13	0.17	0.06
	(0.15)	(0.15)	(0.14)
Log diff. population	0.01	0.04	-0.04
	(0.09)	(0.09)	(0.07)
Change frac. young male	-0.04	-0.04	-0.05
	(0.02)	(0.02)	(0.01)
Change in crimes reported (per 1000 people)	-4.13	-4.65	-3.09
	(15.61)	(17.95)	(9.34)
Observations	465	311	154

Notes: Full sample includes all counties in the eight eastern states for which I have both coal reserves information and total crime data in 1970. The no coal counties are the subset for which there were no known coal reserves. The coal subsample is counties with known coal reserves.

another agency) in 1969. I collapse to the county level (based on primary county, for agencies operating in multiple counties). This leaves me with 9440 county/year observations in the sample states for which I have per capita crime data between 1970 and 1990. Not all counties appear every year due to nonreporting. Estimating effects on changes in the crime rate (as described by equation (2)) uses first differences in per capita crime, so for every year that a county fails to report, I lose two county/year observations–the year of the missing report and the following year. Observations from 1970 to 1972 are also omitted from the sample because I include two lags of the change in the value of coal reserves. The analysis sample for which I have changes in per capita crime as well as the instrument and both its lags has 8,219 county/year observations[6]. The number of counties represented in a year ranges from 392 to 465.

Table 1.1 displays summary statistics for the sample of 465 counties for which I have both coal reserves information and per capita total crime data in 1970, as well as the subsample with no coal reserves and the subsample with coal reserves. The descriptive statistics confirm that counties with coal reserves experienced greater growth in earnings and jobs during the boom, and slower growth during the bust. Systematic differences in demographic trends are also documented in Table 1.1. The share of young men increased slightly more in the boom and decreased by slightly more during the bust for coal counties. Overall population growth in these counties was slightly lower during the boom, and much lower during the bust, when areas with coal experienced population shrinkage. Both subsamples saw rising crime rates during the boom and falling crime rates during the bust. Trends were relatively more favorable for coal counties during the boom and less favorable during the bust.

1.5 Results

1.5.1 Effects on Local Economic Conditions

In Table 1.2, I estimate the effect of the coal boom and bust on log changes in total earnings and total jobs at the county level. The top panel estimates effects for the full sample of counties, and the bottom panel restricts attention to the subset of counties with known coal reserves. Having more coal is associated with higher earnings growth in years when the price of coal rises and one year following the increase. A comparison of the estimates for the full sample and the coal sample indicate that there are substantial nonlinearities in the relationship between the size of coal reserves and local economic impacts: the coefficient estimates for both earnings and jobs are almost an order

[6] I include observations with missing crime data when I estimate earnings and jobs effects. Restricting the sample to county/years for which I observe changes in crime rates produces nearly identical results.

of magnitude larger for the coal sample.

To interpret the coefficients, consider the implied effect size for a county with the mean level of coal reserves, conditional on having coal–about 925 million short tons. In the year of the OPEC embargo (1973 to 1974), the log change in the price of coal was 0.52. Using the estimates from the full sample, this year's spike would have caused earnings in a county with 925 million tons of coal to

Table 1.2: Effects on Earnings and Jobs

	Earnings	Jobs
Full sample		
ln(coalres)*change price	0.002***	0.001
	(0.001)	(0.000)
first lag	0.003***	0.002***
	(0.001)	(0.000)
second lag	0.001	–0.000
	(0.001)	(0.000)
N	11153	11153
Coal counties		
ln(coalres)*change price	0.016***	0.008***
	(0.005)	(0.003)
first lag	0.023***	0.009***
	(0.004)	(0.003)
second lag	–0.007**	–0.004**
	(0.004)	(0.002)
N	3572	3572

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: Each column displays results from a regression where the outcome variable is an annual log change. Column 1 estimates effects on changes in total county earnings, and column 2 estimates effects for changes in the total number of jobs. The regressor of interest is the natural log of coal reserves times the log difference in the real price of coal from year to year and two lags of this variable. The top panel estimates effects for the full sample of counties in states for which I have coal reserves data. The bottom panel restricts attention to counties with known coal reserves. All specifications include controls for log population in 1970, log change in population from year to year, population per square mile in 1970, unemployment in 1970, share of people below the FPL in 1970, and state by year fixed effects. Standard errors are clustered at the county level.

increase 2.1% more than a county with no known coal, plus an additional 3.2% the following year. In contrast, based on estimates from the coal subsample, the 1973 price spike would have boosted earnings in a county with the mean level of coal reserves by 17.2% that year and 24.7% the following year, relative to a county with no coal. Since counties with no known coal are out of sample for the bottom panel, it follows that differences in effect sizes for low-coal counties and no-coal counties are tiny compared with the differences between counties with high levels of coal and low levels of coal. The estimates for the coal subsample are even more striking in the context of the sample averages in Table 1.1. During the boom, the mean growth in earnings during the entire boom decade for coal counties was about 23%. A county with mean coal levels would have come close to doubling this from 1973 to 1975 alone. The coefficient on the second lag is negative and statistically significant for the coal subsample, but the magnitude suggests only a partial reversal of effects over time.

From the second column, it appears that jobs growth may have lagged slightly behind earnings growth. In the top panel, there is a significant effect on the first lag of the treatment variable, but no significant contemporaneous effect of increasing the value of coal reserves. For the coal subsample, the contemporaneous effect and the first lag effect are large, significant, and nearly identical. Both coefficients imply that increases in the value of coal reserves prompted jobs growth that was 8 to 9% higher in a county with the mean level of coal reserves, relative to a county with no reserves, in the year of the OPEC embargo and the following year[7]. The negative coefficient on the second lag again suggests some erosion of effects over time.

1.5.2 Population Effects

Table 1.3 displays estimates of the impact of the mining boom and bust on population characteristics. As described in the methodology section, I cannot estimate effects using a full county/year panel because population data are not collected annually. Instead, I estimate effects on changes during the entire boom decade (1970-1980) in column 1 and the entire bust decade (1980-1990) in column 2. The estimated coefficients in the first row of the top panel suggest a positive (but statistically insignificant) relationship between higher levels of coal reserves and population growth during the boom. During the bust, the relationship is negative, larger in magnitude, and statistically significant. The coefficient indicates that, between 1980 and 1990, a county with the

[7] Black et al. (2002) do not report estimates of the effect of a change in the value of coal reserves on earnings. However, they instrument for changes in earnings using changes in the value of coal reserves in a two-stage model. Their large first-stage F-statistics (>25) show that the value of coal reserves is not a weak instrument for earnings.

Table 1.3: Population Effects

	Boom	Bust
Full sample		
Log difference total population	0.0002	–0.0025***
	(0.0004)	(0.0004)
Change fraction young male	0.0002*	–0.0003***
	(0.0001)	(0.0001)
Change fraction young female	0.0001	–0.0002***
	(0.0001)	(0.0001)
Change fraction older male	–0.0001	0.0002***
	(0.0001)	(0.0001)
Change fraction older female	–0.0002**	0.0003***
	(0.0001)	(0.0001)
Coal counties		
Log difference total population	0.0045*	–0.0041*
	(0.0026)	(0.0022)
Change fraction young male	0.0006	–0.0011*
	(0.0007)	(0.0006)
Change fraction young female	0.0002	–0.0008
	(0.0006)	(0.0005)
Change fraction older male	0.0001	0.0005
	(0.0006)	(0.0005)
Change fraction older female	–0.0008	0.0013**
	(0.0005)	(0.0007)

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: The first row in each panel reports effects on log difference in total population. The first column estimates effects on changes during the boom decade (1970-1980), and the second column estimates effects on changes during the bust decade (1980-1990). The remaining rows estimate effects on the change in the share of the population made up of young males and females (18-30) and older males and females (31-50). The treatment variable is the natural log of coal reserves. All specifications include controls for log population in 1970, population per square mile in 1970, unemployment in 1970, share of people below the FPL in 1970, and state fixed effects. Standard errors are robust. As described in the data section, there are 587 counties in the full sample and 188 counties in the coal sample.

mean level of coal reserves would have had a population growth rate about 5% lower than a county with no coal reserves. The estimated full sample population effects during the bust are large relative to average population growth during this time period (from Table 1.1, about 1%), as well as relative to estimated effects for the boom period, which are an order of magnitude smaller.

The results in the bottom panel once again reveal the nonlinear relationship between the level of coal reserves and the impact of the mining boom and bust. Since coal counties make up a minority of the full sample, the loss of precision in the bottom panel is not surprising. The magnitudes of the coefficients follow the same pattern as in Table 1.2: almost without exception, the estimates for the coal counties are larger than for the full sample. Here the coefficients for total population effects imply that a county with the mean level of coal would have had 9% higher population growth during the boom and 8% lower population growth during the bust than a county with no coal.

The remaining rows of both panels measure population composition effects. As with overall population size, the boom generated smaller compositional changes than the bust. The signs and magnitudes of the estimates suggest the boom led to larger shares of young males, smaller increases in the shares of young females, and shrinkage in the share of older females. In contrast, the bust precipitated big changes in population makeup. Counties with more coal saw their share of young men drop. The drop in the share of young females is smaller, but still at least marginally statistically significant in both panels. The distribution of older individuals (over the age of 30) also shifts to skew more female during the bust. As a result, counties with larger coal reserves are left with older and more female populations at the end of the bust. The smaller boom effects together with the larger bust effects suggest that compositional changes after the price of coal peaked were more powerful than during the ramping up of coal employment in the previous decade. In other words, the collapse of coal prices and mining employment drove a selected group to leave.

1.5.3 Crime Effects

In Tables 1.2 and 1.3, I demonstrated that the mining boom and bust had a substantial impact on the makeup and prosperity of affected areas. Based on the existing literature about economic conditions and criminal behavior, the positive effect of rising coal values on local earnings would be expected to reduce crime rates by pushing down property crime. However, the statistics on crime rates by demographic group suggest that the selective migration generated by the mining boom and bust could have a countervailing effect: young males commit crime at the highest rate,

Table 1.4: Effects on Reported Crimes per 1000 People

	Total	Property	Violent
Full sample			
ln(coalres)*change price	–0.250*	–0.212*	–0.038
	(0.131)	(0.113)	(0.028)
first lag	–0.112	–0.125	0.013
	(0.182)	(0.155)	(0.037)
second lag	0.248	0.232	0.015
	(0.172)	(0.151)	(0.032)
N	8219	8219	8219
Coal counties			
ln(coalres)*change price	–0.386	–0.239	–0.153
	(0.930)	(0.795)	(0.173)
first lag	0.594	0.196	0.403*
	(1.342)	(1.205)	(0.219)
second lag	–0.262	0.048	–0.302*
	(1.512)	(1.393)	(0.169)
N	2777	2777	2777

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: Each column displays results from a regression where the outcome variable is the annual change in crimes per 1000 people. Column 1 estimate effects on changes in the total number of crimes reported, column 2 estimates effects on changes in the number of property crimes reported, and column 3 estimates effects on changes in the number of violent crimes reported. The regressor of interest is the natural log of coal reserves times the log difference in the real price of coal from year to year and two lags of this variable. The top panel estimates effects for the full sample of counties in states for which I have coal reserves data. The bottom panel restricts attention to counties with known coal reserves. All specifications include controls for log population in 1970, log change in population from year to year, population per square mile in 1970, unemployment in 1970, share of people below the FPL in 1970, and state by year fixed effects. Standard errors are clustered at the county level.

and become more prevalent in high coal counties during the boom before leaving during the bust.

In Table 1.4, I explore the net effect on local crime rates. As in Table 1.2, Table 1.4 estimates effects using the year-to-year change in the value of coal reserves. The top panel reports effects based on the full sample, while the bottom table reports effects based on coal counties. Once again, the magnitudes based on coal counties alone tend to be larger, especially for violent crime, although the lack of precision makes inference tentative at best. Both sets of estimates are consistent with a contemporaneous drop in crime as the value of coal reserves increases, driven primarily by changes in property crime rates. The coefficient for total crime using the full sample implies that, for a county with the mean level of coal reserves, reported crimes would have fallen by about 2.7 per 1000 people from 1973 to 1974 relative to a county with no coal. About 85% of this difference comes from reductions in property crime rates. The mean crime rate for counties in my sample between 1970 and 1990 was about 28 crimes per 1000 people. Relative to this average level, a difference in rate change of 2.7 is economically significant. However, considering the cumulative effect across lags suggests the ultimate impact on crime rates was small at best: the coefficient on the second lag imply that the majority of crime rate reductions in the year or two following an increase in the value of coal was reversed. The coefficients on the coal county subsample indicate a statistically significant uptick in violent crime rates on the first lag of the treatment variable, but this effect is mostly undone by the second lag. Overall, the mixed results may reflect offsetting effects driven by sequential changes in local economic conditions and population composition.

1.6 Conclusion

I use the mining boom and bust of the 1970s and 1980s to analyze the effect of a natural resource boom on local crime rates. I find evidence that the boom and bust had a sizeable effect on local conditions that have important theoretical and empirical links to crime rates. Growth in coal prices brought higher growth in jobs and earnings, relative to the years before and after. The rapid and persistent decline in coal prices during the bust ushered in a period of selective population shrinkage, in which affected areas lost relatively more males and young people. Economic prosperity is associated with lower crime rates, while young males are at higher risk for engaging in criminal behavior. Given the large impact of the mining boom and bust on local economic conditions, and the consistent empirical findings that labor market conditions are negatively related to crime, a reasonable ex ante expectation would be that counties with high levels of coal reserves would experience noticeable improvements in crime rates when those reserves become more valuable (and vice versa when the value of coal falls). The crime results are perhaps a cautionary tale for trying to infer how communities will be impacted by large, persistent shocks based on individual-level evidence from transitory shocks. Natural resource booms that lead to substantial inmigration during boom years could be a fruitful avenue for future study, as the introduction of young, low-skill workers without strong community ties may lead to different local dynamics.

CHAPTER 2
THE UNINTENDED IMPACT OF PRETRIAL DETENTION ON CASE OUTCOMES: EVIDENCE FROM NYC ARRAIGNMENTS[1]

2.1 Introduction

On any given day in the United States, over 400,000 individuals (or 1 out of every 550 adults) are in jail awaiting the resolution of their criminal case (Wagner and Sakala, 2014) at a cost of $9 billion per year (Holder, 2011). The number of individuals detained pretrial is also growing, with over a 20% increase between 2000 and 2014 (Wagner, 2015). A large majority of these detainees are prevented from returning home while their cases are adjudicated, because they do not have access to the financial resources to post bail. Of the 38% of felony defendants who are detained, 9 out of 10 of these fail to post bail.[2] (Reaves, 2013) Of those held on bail, 81% have bail set at less than $5,000 and 44% have bail set at less than $1,000 (Phillips, 2012). Given that the high levels of pretrial detention are driven by failure to post bail, even at low levels of bail, low-income individuals are particularly vulnerable to the adverse effects of pretrial detention.

The law allows a judge to set bail or detain a defendant to ensure appearance at court or in the interests of public safety. Pretrial detention is intended to play no further role in legal proceedings. However, detained individuals may face stronger incentives to plead guilty, even if they are innocent. Detainees might miss work and therefore forego income or even lose employment, and they are unable to attend to family responsibilities or access their social support network. For defendants charged with minor offenses, pleading guilty often results in immediate release. Because time spent in jail awaiting the resolution of the case is counted against sentence length, the cost of pleading guilty is lower for detained defendants because they have effectively paid part of the price of conviction in advance.

Empirical exploration of the link between detention and case outcomes began in the early 1960s as part of Vera's Manhattan Bail Project. Ares et al. (1963) compared defendants after controlling for charge type, and found that detention was associated with a higher probability of

---

[1]This chaptered is coauthored with Nolan Pope. The authors would like to thank David Frisvold, Lars Lefgren, Steven Levitt, Jens Ludwig, Justine Olderman, Nathan Petek, Mary Phillips, Joseph Price, Michael Riley, and John Whiston for helpful comments and discussion. We are also grateful for the data provided by the New York State Division of Criminal Justice Services (DCJS). The opinions, findings, and conclusions expressed in this publication are those of the authors and not those of DCJS. Neither New York State nor DCJS assumes liability for its contents or use thereof.

[2]The BJS does not complies national statistics on pretrial detainees in misdemeanor cases.

conviction for all charge types. This relationship was found to be robust to the inclusion of other controls, including criminal record and bail amount (Rankin, 1964). More recently, evidence using multivariate analysis has consistently found a relationship between pretrial detention and case outcomes (Kellough and Wortley (2002); Leiber and Fox (2005); Phillips (2008); Williams (2003)). Although these results are certainly consistent with a causal effect of detention on conviction, they cannot rule out the possibility that systematic, unobserved differences are responsible for part or all of the relationship.

In this paper, we instrument for pretrial detention status using the degree to which individual judge-specific detention rates deviate from the average detention rate for the full sample by crime type, allowing us to isolate the causal effect of detention on conviction and other case outcomes. We use data from nearly a million felony and misdemeanor criminal cases in New York City from 2009 to 2013, including information linking defendants with arraignment judges. Once we confirm the conditional random assignment of judges, we estimate how an increase in the probability of pretrial detention, driven solely by the arraignment judge, affects the probability of pleading guilty, being convicted, and sentence length.

We find that pretrial detention increases the probability that a felony defendant will be convicted by at least 13 percentage points; effects for misdemeanor defendants are larger than seven percentage points. The increase in conviction rates is driven by detainees accepting plea deals more frequently. We also find evidence that detention increases minimum sentence length. Detention affects case outcomes for defendants in both misdemeanor and felony cases. In addition, individuals who are detained pretrial are less likely to obtain a reduction in the severity of the crime they are charged with. Our estimates are robust to adding a variety of demographic, criminal history, and most serious offense charge controls. Relaxing the monotonicity assumption associated with standard instrumental variables (IV) also yields similar results. Although effects do not vary noticeably across racial and ethnic lines, minorities are overrepresented in the sample and are therefore disproportionately affected by pretrial detention.

The adverse effects of detention on case outcomes could operate through a variety of mechanisms. Conventional wisdom among those employed in the criminal justice system is that detainees plead guilty to return home quickly. Large effects for people charged with minor crimes support this hypothesis, because many of these individuals have access to plea deals that do not require any additional time incarcerated. Defendants with dependents may be especially prone to pleading guilty quickly so they can be with their children. We find evidence that women are more likely than men to plead guilty when they are facing less serious charges; we find no difference in effects for men and women facing more serious charges. However, even individuals who will almost certainly spend more time behind bars after pleading guilty tend to do so more often if detained. This group has already served part of their time in advance, making accepting a plea deal relatively less costly.

One existing branch of literature uses judge assignment to analyze post-sentence outcomes for convicted criminals: Kling (2006) finds no consistent evidence of adverse effects from serving a longer sentence on employment and earnings. Aizer and Doyle (2015) find that juvenile incarceration decreases the likelihood of high school completion and increases the likelihood of adult recidivism. Mueller-Smith (2015) finds that a one-year prison term generates \$56,200 to \$66,800 in social costs through increased recidivism, decreased employment and wages, and increased dependence on public assistance.

Concurrent work has taken a similar approach to exploring the impact of pretrial detention. Stevenson (2016) uses variation in the bail-setting patterns of arraignment magistrates in Philadelphia to find that defendants who fail to make bail are more likely to be convicted. Gupta et al. (2016) also use data from Pennsylvania, and estimate that being assigned money bail increases the probability of both pleading guilty and reoffending. Dobbie et al. (2016) use data from both Philadelphia and Miami and find that pretrial detention increases convictions and employment. By contrast, our data from New York City, which has the second-largest jail system in the country, allows us to estimate effects with greater precision using a much larger sample of arraignments and arraignment judges.

The paper proceeds as follows. Section 2 describes the institutional setting of the New York City arraignment process. Section 3 describes the two data sets used in the analysis. Section 4 describes the IV methodology and discusses potential concerns with it. Section 5 presents the main results from our analysis along with suggestive evidence for the underlying mechanism and robustness checks. Section 6 concludes.

2.2 Background on New York City Arraignment Process

In New York City, after an individual is arrested, he is taken to the local police precinct to be booked and fingerprinted. Prosecutors are assigned immediately after arrest, and are responsible for the case until disposition by trial or plea. Prosecutor assignment is not linked to judge assignment, unlike some jurisdictions. The fingerprints and booking information are processed by the state and are used to provide a criminal history of the defendant for the arraignment judge. Individuals are then moved to a holding cell in the county courthouse corresponding to the county of arrest (Bronx, Kings, New York, Queens, or Richmond) to await their arraignment. A small subset of misdemeanor arrestees are given desk-appearance tickets and are not held prior to arraignment. At the courthouse, the Criminal Justice Agency will interview the defendant and will provide a bail recommendation for the arraignment judge. The defendant is then allowed to meet with his defense attorney, which may be private or assigned, before the arraignment.

The process of judge assignment is central to our identification strategy. If individuals are sorted among arraignment judges based on unobservables that are correlated with case outcomes (e.g. the strength of evidence against them), then judge-level variation in propensity to detain (through denial of bail or systematically setting higher bail amounts) will not be a valid instrument. Arraignments in New York City happen every day of the year in two sessions: day (9am to 5pm) and night (5pm to 1am). In a session, judges only have an average of 6 minutes per arraignment. Most felony defendants are arraigned within one to two days of their arrest, during which time they are kept in a holding cell.[3] Starting before our sample period, the New York City criminal courts received a mandate from the Court of Appeals to arraign within 24 hours. The average time from arrest to arraignment in our sample period fell from 25.39 hours in 2009 to 21.44 hours in 2013 (Barry, 2014). Given the short window between arrest and arraignment, police officers are unlikely to manipulate which judge a defendant will see at arraignment. Arraignments are scheduled by Arraignment Coordinators, whose primary objective if multiple courtrooms are operating is maintaining balance in the workloads across courtrooms. Some intentional sorting across arraignment shifts does occur. For example, defendants with desk-appearance tickets, who are generally charged with relatively minor offenses, are grouped together. Thus, there is no institutional reason to expect troublesome sorting of defendants among judges conditional on the courthouse, time of year, and arraignment session. Below, we present further statistical evidence that judge assignment is conditionally random.

The arraignment is the first time the defendant will appear before a judge. Here, he is formally informed of the charges being brought against him. In felony cases, the judge will release the defendant on his own recognizance, set a bail amount, or order the defendant held without bail (Phillips, 2012).[4] Misdemeanor cases may be adjudicated at arraignment if the defendant pleads guilty or the case is dismissed. In practice, defendants are generally not allowed to plead guilty

[3] Individuals facing a charge in the least serious category, nonviolent class E felonies, may be released awaiting an arraignment. In this situation, the arraignment is scheduled weeks or months ahead of time.

[4] With the exception of Queens, supervised release is not available in New York City.

at arraignment if there is a complaining witness (e.g., the assault victim in an assault case). In determining the terms of a guilty plea at arraignment, the judge may not offer reduced charges, but can offer specific punishments for the existing charges. There is less room for judicial discretion in determining whether a case will be dismissed, because judges may dismiss a case at arraignment only when a specific legal defect exists. Because some misdemeanor cases are disposed of at arraignment, and we do not observe which of these defendants would have been detained if their cases had not been immediately resolved, we interpret estimates from the misdemeanor subsample with more caution than those from the felony subsample.

In New York county courts, aside from being prohibited from denying bail in misdemeanor cases or for the sake of public safety, the judge has complete discretion in the bail decision (CPL 530.20). Judges may set different amounts for different types of bail; for example, a defendant could be required to post either \$1000 in cash, or secure a \$2000 bond. Following Phillips (2012), we use the amount of cash required for release. If the cash alternative is not set, we use the bond amount. Defendants who fail to make bail are detained until their cases are adjudicated. In New York City, almost all pretrial detainees are housed in jails on Rikers Island, with the remainder kept in county jails throughout the city. Cases may be resolved at trial, with a conviction or acquittal, but they may also be resolved through plea deals and dismissals. Trials are far less common than plea deals; only 0.7% of cases in our sample go to trial, whereas 64.3% are disposed of through plea deals. A large majority of the remaining cases have the charges dismissed. The trial judge is assigned separately from the arraignment judge. A different judge is randomly assigned to the case after the arraignment, so that the arraignment judge can only influence case outcomes through pretrial-status assignment (Uniform Rules for Trial Courts [22 NYCRR] 200.11(c)). We are interested in how the detention status resulting from the arraignment hearing affects case outcomes, including guilty verdicts, guilty pleas, the relationship between arraignment charges and disposition charges, and sentencing.

2.3 Data

The data come from two sources: the State Court Processing Statistics (SCPS) and the New York City Criminal Courts (NYC). The SCPS data include a random sample of felony cases in 40 of the 75 most populous counties in the United States for even years from 1990 to 2006. In addition to information on pretrial status and case outcomes, we take advantage of demographic (age, race, and sex) and criminal history (prior arrests and convictions) details. We supplement the SCPS data with county-level data from the Census Bureau on crimes reported per capita and median income. The sample is restricted to individuals for whom we observe both pretrial status and the final case outcome. This eliminates 3,121 cases with no pretrial-status information, 12,980 cases that were still pending at the time of data reporting, 6,629 cases that were diverted or deferred, and 237

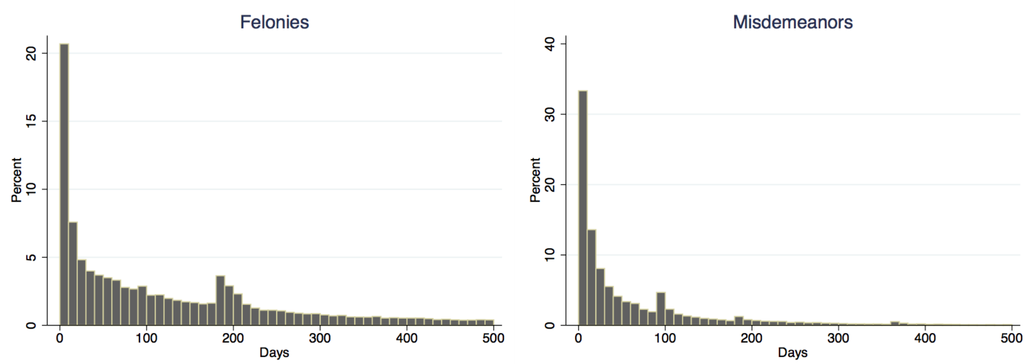
Table 2.1: SCPS Summary Statistics

	Full Sample	Released	Detained
Demographics			
Age	30.4	30.2	30.8
Female	0.17	0.20	0.12
White	0.29	0.33	0.24
Black	0.44	0.44	0.44
Hispanic	0.25	0.21	0.30
Criminal History			
First Time Offender	0.39	0.50	0.26
Prior Felony Arrests	3.1	2.3	4.0
Prior Felony Convictions	1.3	0.9	1.7
Prior Misdemeanor Arrests	3.2	2.6	3.9
Prior Misdemeanor Convictions	1.8	1.3	2.3
Arraignment Information			
Detained Pretrial	0.46	0.00	1.00
Released on Own Recognizance	0.27	0.50	0.00
Released on Bail	0.26	0.49	0.00
Held on Bail	0.35	0.00	0.76
Remanded	0.07	0.00	0.15
Case Outcomes			
Convicted	0.71	0.65	0.79
Plead Guilty	0.67	0.62	0.74
Guilty at Trial	0.04	0.04	0.05
Sentenced to Incarceration	0.49	0.34	0.66
Max Sentence Length (Days)	1055	241	1995
Dismissed	0.25	0.31	0.19
Observations	82,114	43,975	38,139

Notes: This table contains summary statistics of all arraignments in the SCPS sample. The first column is for the entire sample. The second column is for individuals released any time after their arraignment. The third column is for individuals detained between their arraignment and the resolution of their cases.

cases with no case-outcome information. An additional 168 observations are missing criminal-history information (prior arrests and/or convictions), and 13,336 have no information about race/origin.[5] This leaves us with a final sample size of 82,114 Table 2.1 presents summary statistics for the SCPS data. The disparity in outcomes by pretrial status is stark: 79% of detained individuals are convicted, compared to 65% of those who are released. However, the systematic differences in other observables

Figure 2.1: Length of Detention if Detained



Notes: Each graph is a histogram of the number of days an individual is detained conditional on being detained. For each bar, the y-axis shows what percent of individuals were detained for the indicated number of days. Each bar has a width of 10 days. The first graph is for felonies and the second is for misdemeanors.

are just as extreme; in particular, detained individuals have many more prior arrests and convictions. Although controlling for observable characteristics may yield a more believable estimate of the effect of detention on conviction, it cannot address the obvious concern that two groups that are systematically different in observable ways are likely to be different in unobserved ways as well. Thus, without a quasi-experimental approach, using exogenous variation in the probability of pretrial detention, there is good reason to believe our estimates will be biased.

The NYC data include additional arraignment details that allow us to take advantage of exactly this kind of natural experiment. Our data cover all felony and misdemeanor criminal cases

[5]The results are not sensitive to omitting race/origin controls from the analysis and including these observations

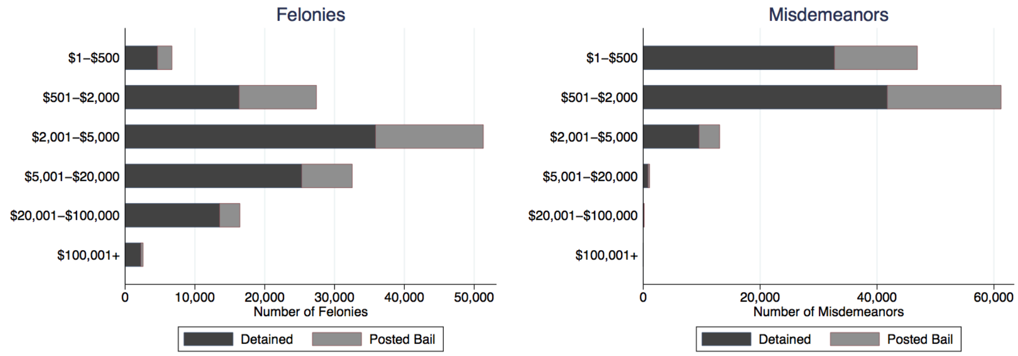
Table 2.2: NYC Summary Statistics

	Felonies		Misdemeanors		All
	Released	Detained	Released	Detained	
Demographics					
Age	31.0	32.5	31.6	35.4	31.9
Female	0.20	0.10	0.19	0.13	0.17
White	0.13	0.08	0.14	0.10	0.13
Black	0.46	0.56	0.46	0.59	0.48
Hispanic	0.35	0.34	0.34	0.29	0.34
Criminal History					
First Time Offender	0.68	0.36	0.66	0.24	0.59
Prior Felony Arrests	2.2	5.5	2.2	6.8	2.9
Prior Misdemeanor Arrests	3.0	7.6	3.8	13.2	5.0
Prior Felony Convictions	0.4	1.2	0.4	1.3	0.6
Prior Misdemeanor Convictions	1.4	5.2	1.8	10.1	2.9
Sex Offender	0.01	0.02	0.01	0.02	0.01
Arraignment Information					
Detained Pretrial	0.00	1.00	0.00	1.00	0.20
Released on Own Recognizance	0.72	0.00	0.94	0.00	0.72
Released on Bail	0.28	0.00	0.06	0.00	0.08
Held on Bail	0.00	0.93	0.00	0.95	0.19
Remanded	0.00	0.07	0.00	0.05	0.01
Number of Counts	1.4	1.2	1.1	1.1	1.1
Case Outcomes					
Convicted	0.64	0.79	0.65	0.81	0.68
Plead Guilty	0.59	0.70	0.63	0.77	0.64
Dismissed	0.33	0.18	0.34	0.19	0.31
Went to Trial	0.01	0.02	0.00	0.00	0.01
Sentenced to Incarceration	0.11	0.53	0.04	0.50	0.15
Charge Reduction	0.84	0.72	0.79	0.33	0.74
Min Sentence Length (Days)	71	400	15	35	66
Max Sentence Length (Days)	85	532	15	37	83
Observations	139,801	105,259	639,141	89,614	973,815

Notes: This table contains summary statistics of all arraignments in the NYC DCJS sample. The first two columns are for felonies, the third and fourth columns are for misdemeanors, and the last column is for the entire sample. The first and third columns are for individuals released any time after their arraignment. The second and fourth columns are for individuals detained between their arraignment and the resolution of their cases.

between 2009 and 2013. To have a precise measure of judge severity, we restrict cases to those in which the arraignment was brought before a judge who saw at least 500 arraignments in our sample. This eliminates 24,679 cases (2.5% of the original sample). Along with pretrial status, bail amount, demographics, criminal history, and case outcomes, the data include the arraignment shift,

Figure 2.2: Bail Amount



Notes: Each graph shows the number of defendants for each bail amount. For each bar, the y-axis shows the bail amount and the x-axis shows the number of defendants at the given bail amount. The first part of each bar shows the fraction of defendants who were detained pretrial. The second part of each bar shows the fraction of defendants who posted bail and were released pretrial.

courthouse, and judge ID. Arraignments in our sample were presided over by 212 unique judges, all of whom presided over arraignments for both felony and misdemeanor cases. The average number of arraignments per judge per year is 1,523. Of these, 291 are felony cases. We treat a defendant as detained pretrial if he was remanded without bail, or bail was set and never posted.[6] As a result, individuals who posted bail after spending some time in jail are counted as released, which could bias our results toward zero. Figure 2.1 presents graphically the distribution of pretrial detention length. The bumps in the distributions around 180 days for felonies and 90 days for misdemeanors are a result of laws meant to protect the right to a speedy trial: the state is generally required to bring cases to trial within six months for felony cases or 90 day for misdemeanor cases, although prosecutors can successfully move for an extension in a variety of circumstances (CPL 30.30). Appendix Figure B.1 shows how the sample is distributed across the most common types of offenses. Table 2.2 displays summary statistics for the NYC sample. In many ways, the NYC felony sample and national SCPS sample are comparable. In terms of pretrial treatment, the fraction released pretrial is similar, but being released on bail is much less common in the NYC data than it is in the SCPS data (27% of

[6] We observe whether, but not when, bail was posted, and therefore have no way of determining how many defendants in our sample are detained for the full period between arraignment and case disposition.

those released vs. 49%). This suggests the margin between being released on one's own recognizance and having bail set at all is more important in New York City than nationally, whereas, at least for the period of our SCPS data, a larger fraction of individuals facing bail are able to come up with enough money to secure their release. Figure 2.2 shows the distribution of bail amounts and illustrates that, regardless of the bail amount, defendants are more likely to be held on bail than to post bail. The summary statistics in Table 2.2 also make clear that, with few exceptions, cases either end in conviction or dismissal. Thus, although we do not estimate effects on dismissal directly, the impact of detention on dismissal would be equal and opposite to the estimates we find for conviction.

2.4 Methodology

Our baseline OLS results come from estimating the following equation:

$$y_i = \beta_o + \beta_1 detained_i + \beta_2 X_i + \beta_3 C_i + \beta_4 T_i + \epsilon_i,$$

where y_i is a dummy for whether or not defendant i was convicted, a dummy for whether or not he pled guilty, or the minimum sentence length. The variable *detained* is a dummy for being detained pretrial. X_i is a vector of demographic characteristics, including race, a cubic in age, gender, and a police precinct fixed effect. C_i is a vector of criminal-history features, including dummies for first-time offender, sex-offender status, whether the current charge is an attempt charge, whether it is a violent felony, whether a firearm was involved, whether a weapon was involved, whether a child was involved, whether it is a hate crime, whether it is a drug crime, the class of the most serious offense charge (e.g., class B felony), Uniform Crime Reporting crime-type code for the most serious offense charge (e.g., aggravated assault, DUI), and cubics in the number of prior felony arrests, misdemeanor arrests, felony convictions, misdemeanor convictions, and the number of counts in the current case. T_i is a vector of fixed effects for year, month, day of week, and courtroom/shift (e.g., Tuesday night shift in courtroom A of the Bronx courthouse in July 2012). Standard errors are clustered at the arraignment-shift level, which is more conservative than clustering at the judge-year or courthouse-month level.

Our objective is to estimate the causal effect of being detained on case outcomes, which we do by instrumenting for *detained* in the above regression equation using judge-level variation in the propensity to detain pretrial for a given crime type.[7] Specifically, we calculate a leave-out mean

[7] In theory, using judge fixed effects as our instruments would introduce bias. In practice, we find the results are qualitatively similar in our context (see Tables B.4 and B.5 in the appendix).

measuring the degree to which each judge deviates from the full-sample average detention rate for each Uniform Crime Reporting crime type code:

$$z_{ijc} = \frac{1}{n_{jc}-1}(\Sigma_k \widehat{detained}_k - \widehat{detained}_i) - \frac{1}{n_c - 1}(\Sigma_m \widehat{detained}_m - \widehat{detained}_i),$$

where the first term divides the total number of people charged with crime type c and detained by judge j (indexed by k and excluding individual i) by the total number of defendants charged with crime type c and who appeared before judge j (n_{jc}), excluding individual i; and the second term is the total number of people in our sample who were charged with crime type c and detained (indexed by m and excluding individual i), divided by the total number of people charged with crime type c in the sample (n_c), excluding individual i. To control for systematic differences across shifts, we use $\widehat{detained}$, the residual from regressing *detained* on courtroom/shift fixed effects. Positive values of z_{ijc} correspond with judges that detain more than average for the sample, and negative values indicate lower-than-average detention rates.

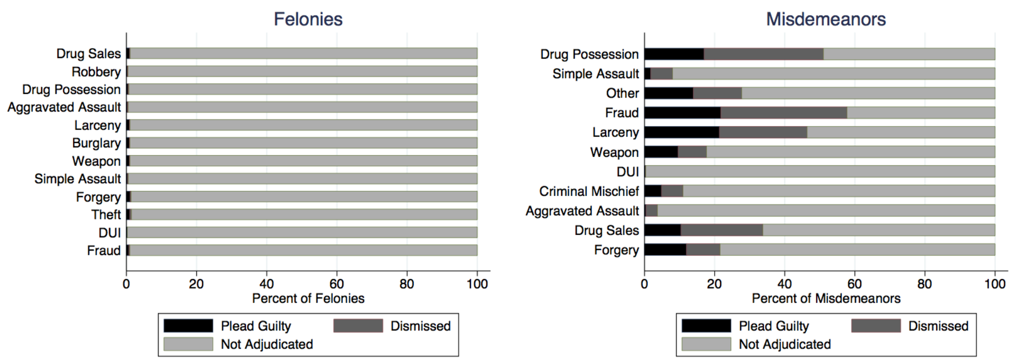
One concern with standard IV in this case is that the monotonicity assumption may be violated. This assumption requires that, if defendants on average are more likely to be detained by judge A than judge B, anyone who would be detained by judge B would also be detained by judge A (Imbens and Angrist, 1994). This assumption could be violated if arraignment judges vary their treatment based on defendant characteristics.[8] To relax the monotonicity assumption, we follow Mueller-Smith (2015) in constructing a set of candidate instruments that measure the degree to which judges deviate from average trends across the full sample in their tendency to detain people within certain groups. These groups are defined by criminal-history features (a dummy for first-time offender and sex offender, and the number of prior felony arrests, prior misdemeanor arrests, prior felony convictions, prior misdemeanor convictions, and the number of counts in the current case), most serious charge (Uniform Crime Reporting crime-type code), race, and gender, as well as interactions between every pairwise combination of these characteristics. We use a Lasso procedure described in Belloni et al. (2012) to reduce noise and avoid the many-instruments problem by selecting only the most informative of these instruments. This process leaves us with 27 instruments for the felony subsample and 23 for the misdemeanor subsample; the five most powerful instruments for each sample are listed in Appendix Table B.1. The new, less restrictive, monotonicity assumption

[8] See the online appendix of Mueller-Smith (2015)for a detailed discussion of how the monotonicity assumption could be violated, and an example illustrating how this violation could bias the results in a similar context.

for these specifications is that judge detention decisions are monotonic within groups of defendants sharing similar characteristics.

As mentioned above, analysis based on misdemeanor defendants should be viewed with some caution. Unlike felony cases, misdemeanor cases may be disposed of at the arraignment, if the case is either dismissed or the defendant chooses to plead guilty. The fraction adjudicated at arraignment by crime type is shown in Figure 2.3. As discussed in more detail below, the results are somewhat sensitive to the inclusion or exclusion of cases that are adjudicated at arraignment.

Figure 2.3: Fraction Who Plead Guilty and Dismissed at Arraignment by Crime Type



Notes: Each graph shows the percent of defendants who plead guilty, had their case dismissed, or did not have their case adjudicated at the arraignment for different crime types. For each bar, the x-axis shows the percent of arraignments that were adjudicated and the y-axis shows the crime type. The first part of each bar shows the percent of arraignments in which the defendant plead guilty at the arraignment. The second part of each bar shows the percent of arraignments in which the case was dismissed at the arraignment. The third part of each bar shows the percent of arraignments that were not adjudicated at the arraignment.

Table 2.3: Test for Judge Random Assignment

	Felonies		Misdemeanors	
	Predicted	Actual	Predicted	Actual
Detained	1.44	10.36	1.26	19.29
Released on Own Recognizance	1.39	19.80	5.46	18.35
Set Bail	1.39	19.18	1.54	30.90
Remanded	1.46	3.22	1.49	6.63

Notes: Each cell in the table reports the F statistic when testing for the joint significance of the judge binary variables for the indicated outcome variable conditional on courthouse by time of year by arraignment shift fixed effects. For the first and third columns, the outcome variables are the predicted values for the indicated variables in the variable column that are obtained from a regression of the indicated variable on demographic, criminal history, and courthouse by time of year by arraignment shift fixed effects controls. The second and fourth columns test for the joint significance of the judge binary variables when the actual (not predicted) indicated variables are the outcome variables.

The randomness of arraignment-judge assignment is essential to the validity of our estimation strategy. To test for conditionally random judge assignment, we first estimate the ex ante probability that each defendant would be detained based on his demographic and criminal-history characteristics, as well as courthouse by time of year (year, month, and day of the week) by arraignment-shift fixed effects (hereafter referred to as time by arraignment-shift fixed effects). We then regress these predicted probabilities on a full set of judge dummy variables controlling for courthouse by time of year by arraignment-shift fixed effects, and test for the joint significance of the judge effects. For example, these regressions test for the random assignment of judges within morning arraignment shifts at the Queens courthouse on Tuesdays in July of 2012. Table 2.3 compares the F-statistics from these regressions with the F-statistics from testing the joint significance of judge effects in a regression on *actual* pretrial detention status controlling for time by arraignment-shift fixed effects. The F-statistics from the tests using predicted pretrial status probabilities are generally less than 2. These levels may represent a technical rejection of the null hypothesis, but reflect relatively small differences in the ex ante characteristics of defendants that appear before different judges. By contrast, the tests using actual pretrial status yield high F-statistics, confirming that judge assignment does in fact matter substantially in terms of a defendant's arraignment outcome. Appendix Tables B.2 and B.3 present further evidence supporting the conditionally random assignment of arraignment judges. Table B.2 reports F-statistics from a covariate-by-covariate analysis of the importance of judge fixed effects. The results show that judge fixed effects are poor predictors of observable demographic and criminal-history characteristics, but are good predictors of pretrial-status assignment. In Table B.3, we regress each predicted and actual pretrial status dummy on the normalized leave-out mean for how much a judge deviates from the full-sample average detention rate for each crime type. The table displays coefficients on the normalized leave-out mean for each of these regressions. Although some of the coefficients from the predicted status regressions are statistically significant, they are small – about a tenth the size of the coefficients on the corresponding actual pretrial status outcomes.[9] Thus, arraignment judges see fairly comparable groups of defendants ex ante, but their arraignment outcomes vary substantially across judges.

2.5 Results

2.5.1 OLS on a National Sample

We begin by analyzing the national SCPS sample. Table 2.4 presents the results of a series of OLS regressions. Even after controlling for county, year, most serious offense charge, sex, age, race, and criminal-history features, being detained pretrial still has a positive and statistically significant relationship with case outcomes. According to our most conservative estimates, being detained is associated with an 8.7-percentage-point increase in the probability of conviction, a 7.4-percentage-point increase in the probability of pleading guilty, and a maximum sentence length just over two and a half years (922 days) longer. Estimates using the same specification when the sample is restricted to cases from New York City are noticeably larger: being detained is associated with a 16.2-percentage-point increase in the probability of conviction, a 15.3-percentage-point increase in the probability of pleading guilty, and a maximum sentence length about 4.5 years longer.

Although simple OLS yields a useful quantitative description of the correlation between detention and case outcomes, it cannot be relied on for valid causal inference unless one is willing to assume pretrial status is uncorrelated with unobservables that also affect the case outcome. This

[9] In all three tables, judge fixed effects are less powerful as predictors of whether or not a defendant was remanded, i.e. detained unconditionally without the opportunity to post bail. This may reflect the fact that the only legally permissible case in which bail can be denied is when no amount would be sufficient to ensure the defendant's return to court(Phillips, 2012).

Table 2.4: OLS Specifications for SCPS Data

	Felonies			
	(1)	(2)	(3)	(4)
Panel A: Conviction				
Detained	0.136***	0.097***	0.097***	0.086***
	(0.003)	(0.003)	(0.003)	(0.003)
Panel B: Plead Guilty				
Detained	0.121***	0.082***	0.083***	0.073***
	(0.003)	(0.003)	(0.003)	(0.003)
Panel C: Maximum Sentence Length (Days)				
Detained	1769***	1083***	1018***	903***
	(118.3)	(93.4)	(92.3)	(96.9)
County, Year, Offense FE		X	X	X
Demographics			X	X
Criminal History				X
Observations	81,332	81,332	81,332	81,332

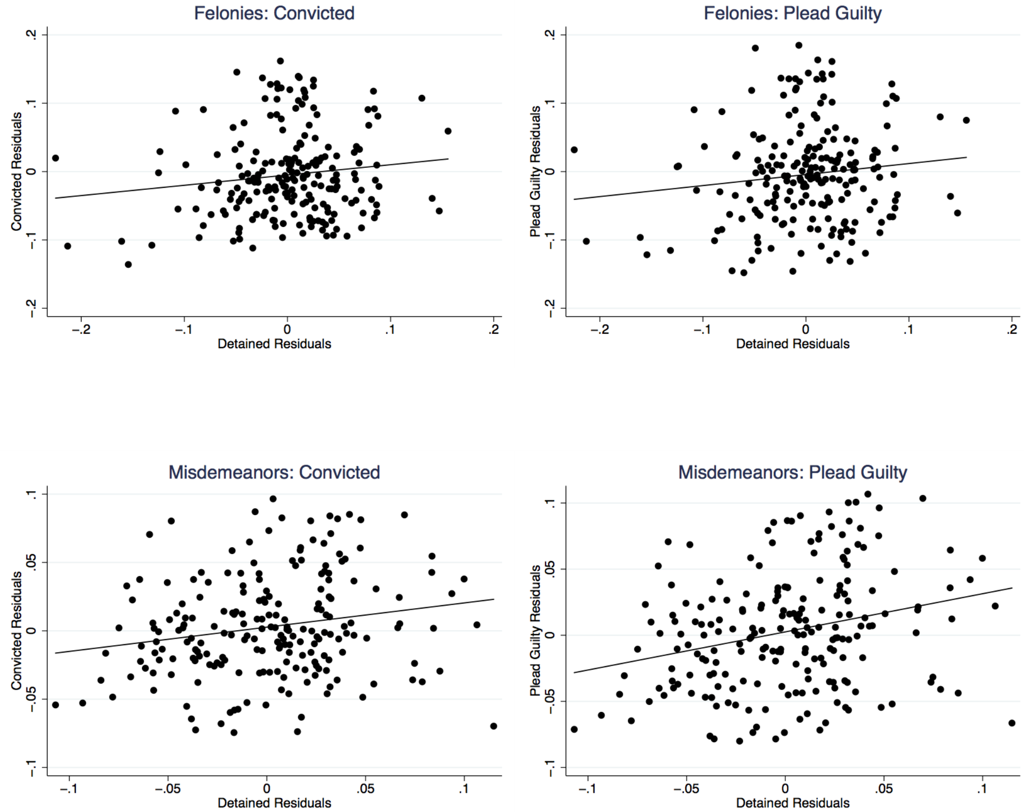
Notes: This table shows estimates of the coefficient on the binary variable detained when the indicated outcome variable is regressed on detained and the indicated controls for the SCPS data. Column 1 contains no controls. Column 2 includes county, year, and offense fixed effects. Column 3 adds demographic controls. Column 4 adds criminal history controls. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

assumption could be violated in several ways. For example, if the family and friends of a defendant are more willing to help provide financial support for bail if they believe him to be innocent, and their beliefs are correlated with actual guilt or innocence, the identifying assumption would be violated. If judges receive some signal of actual guilt or criminality that is unobservable to us, and use the signal in their assignment of pretrial status, this would also generate correlation without causation. Because the SCPS do not include arraignment-judge IDs, we turn our attention to the rich data provided by the New York Criminal Courts for our IV analysis.

2.5.2 IV Results

Figure 2.4 displays graphical evidence of the causal relationship between detention and case outcomes. These graphs plot judge-specific outcome residuals by judge-specific detention residuals.

Figure 2.4: IV Scatterplots



Notes: Each graph plots judge-specific outcome residuals by judge-specific detention residuals. The y-axis represents conviction or plead guilty residuals created by regressing the indicated outcome variable on demographic, criminal history, and time and courtroom fixed effects controls. Similarly, the x-axis represents detention residuals created by regressing the detained variable on demographic, criminal history, and time and courtroom fixed effects controls. Each point of the graph represents the average outcome residual and detention residual for a given judge. The black line is the fitted linear regression through these points.

Each point represents the average outcome residual and detention residual for an individual judge. Average outcome residuals reflect the ultimate outcomes of cases that appeared before the judge at arraignment (not those over which the judge presided at trial). The slope of the fitted linear regression lines confirms that simply appearing before a harsher arraignment judge increases a defendant's likelihood of being convicted and pleading guilty.

These effects are estimated formally for our sample of felony defendants in the first three

Table 2.5: OLS Specifications for NYC Data

	Felonies				Misdemeanors			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Conviction								
Detained	0.144***	0.150***	0.149***	0.149***	0.152***	0.161***	0.143***	0.100***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Panel B: Plead Guilty								
Detained	0.115***	0.120***	0.118***	0.126***	0.147***	0.159***	0.139***	0.096***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Panel C: Minimum Sentence Length (Days)								
Detained	330***	330***	315***	159***	19***	18***	17***	11***
	(3.738)	(3.900)	(3.848)	(3.060)	(0.376)	(0.384)	(0.386)	(0.427)
Time and Courtroom FE		X	X	X		X	X	X
Demographics			X	X			X	X
Criminal History				X				X
Observations	245,060				728,755			

Notes: Table 2.5 shows estimates of the coefficient on the binary variable detained when the indicated outcome variable is regressed on detained and the indicated controls. Columns 1 and 5 contain no controls. Columns 2 and 6 include time and courtroom fixed effects. Columns 3 and 7 add demographic controls. Columns 4 and 8 add criminal history controls. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table 2.6: Effect of Detention on Felony Case Outcomes Using IV Specifications

	Felonies					
	IV			IVNM		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conviction						
Detained	0.109***	0.124***	0.142***	0.162***	0.154***	0.130***
	(0.020)	(0.021)	(0.020)	(0.005)	(0.005)	(0.006)
Panel B: Plead Guilty						
Detained	0.080***	0.093***	0.113***	0.138***	0.129***	0.102***
	(0.021)	(0.021)	(0.021)	(0.005)	(0.005)	(0.006)
Panel C: Minimum Sentence Length (Days)						
Detained	220***	196***	73***	251***	235***	157***
	(28)	(29)	(26)	(7)	(7)	(8)
Time and Courtroom FE	X	X	X	X	X	X
Demographics		X	X		X	X
Criminal History			X			X
F Statistic	2,136	2,052	2,195	1,078	1,064	1,176
Number of Instruments	1	1	1	27	27	27
Observations	243,375	243,375	243,375	242,241	242,241	242,241

Notes: This table reports estimates of the effect of pretrial detention on case outcomes for felonies using instrumental variables. The first three columns use the judge-level variation in the propensity to detain pretrial for a given crime type as an instrument for being detained with varying controls. The last three columns use the set of Lasso selected instruments from the full set of constructed candidate instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

columns of Table 2.6, which show the results from a traditional IV specification using the degree to which each judge deviates from the full-sample detention rate by crime type as an instrument for whether an individual was detained at arraignment. Table 2.5 presents the corresponding OLS estimates for comparison. These estimates are larger than the coefficients from the national sample, but comparable to OLS results using only the subset of SCPS observations from counties in New York City. The IV estimates on the felony subsample in Panel A are slightly smaller than the OLS estimates, although adding more controls attenuates this difference. The estimated effects are significant, both statistically and economically: pretrial detention increases the probability of conviction by over 10 percentage points. The estimated effect of detention on pleading guilty is only about 20% smaller than the effect on conviction (11 percentage points vs. 14 percentage points, using the full set of controls), suggesting that detention primarily affects conviction by inducing some individuals who would not have pled guilty if released to plead guilty after they are detained.

The marginal defendant in this context is one who does not have access to the resources required to make bail. Although judges may differ in both the frequency with which they set bail and the level at which they set it, the decision to set it or not appears to be more important in this setting. In our sample, 71% of felony defendants and 68% of misdemeanor defendants for whom bail is set are unable to post bail to secure their release. Most of the bail amounts standing between these defendants and their release are quite low: around 15% of defendants in felony cases that are held on bail would need less than \$2000, and the majority have bail set at less than \$5000. For the misdemeanor sample, well over half of individuals held on bail need less than \$2000, and over a fourth need only a few hundred dollars. These statistics highlight the unfortunate reality that the unintended adverse effects of pretrial detention on case outcomes fall most heavily on the shoulders of the poor and disadvantaged.

We also present estimates from specifications relaxing the monotonicity assumption for the felony subsample in columns 4-6 of Table 2.6 (labeled "IVNM" for IV non-monotonicity). This correction produces coefficients that are quite similar to the standard IV estimates. Across all the specifications in Table 2.6, we see consistent evidence that being detained pretrial causes the probability of conviction to increase substantially, and that most–if not all–of this increase comes through detainees' guilty pleas. This finding is robust to a variety of estimation methods and controls. In our preferred specification, in which we relax the monotonicity assumption and include the full set of controls, being detained increases the probability of conviction by 13 percentage points and the probability of pleading guilty by 10 percentage points.

Panel C shows evidence that the impact of detention extends to minimum sentence length[10]. The coefficients are less stable, but all suggest a strong, positive effect of being detained. Because we are estimating effects for the full sample, including people who were not convicted, these estimates reflect the combined effect on conviction (a precondition to receiving a sentence) and sentence length once convicted. Our preferred specification suggests that being detained pretrial increases minimum sentence length for felony defendants by over 150 days.

2.5.3 Misdemeanor Cases

Because of the larger variety of outcomes possible at a misdemeanor arraignment, we are more cautious about our analysis of this subsample. Columns 5-8 of Table 2.5 show OLS estimates of the effect of detention on conviction for people charged with misdemeanors. Overall, the estimates are similar to those from the felony subsample. The IV estimates for this group are more responsive to additional controls, but the standard IV and IVNM specifications yield similar results (see Table 2.7). Our preferred specification indicates that being detained increases the probability of conviction by 7.4 percentage points, and the probability of pleading guilty by 7.1 percentage points. In Table 2.6, we show results from a subset of our specifications using various sample constructions that include or exclude cases that were dismissed or in which the defendant pled guilty at arraignment. Both the IV and the IVNM coefficients are sensitive to the sample construction, though in different ways. Despite this sensitivity, almost all specifications indicate a positive and statistically significant effect of detention on the probability of conviction and pleading guilty. As an additional check, we analyze a subset of misdemeanor cases that are the least likely to be adjudicated at arraignment. Figure 2.3 shows the fraction of cases adjudicated at arraignment by most serious offense charge. Simple assault, aggravated assault, and DUI cases are the least likely to be disposed of at the arraignment hearing, especially through a guilty plea. Appendix Table B.7 shows results from running our analysis on this subset of misdemeanor cases. Coefficients for crimes with low levels of adjudication at arraignment are substantially higher than for the full misdemeanor sample. Of course, the defendants in this group are a nonrandom sample, but the large effects do provide evidence that the sample-selection issues created by arraignment adjudications are not generating the large and statistically significant effects we estimate for the misdemeanor sample. Thus, although we are less confident in magnitude of the effect for misdemeanors, we find strong evidence that pretrial detention influences case outcomes

[10] These estimates are not directly comparable to the sentence-length results on the national SCPS sample in Table 2.4, because the SCPS data only include information on *maximum* sentence length.

Table 2.7: Effect of Detention on Misdemeanor Case Outcomes Using IV Specifications

	Misdemeanors					
	IV			IVNM		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conviction						
Detained	0.192***	0.179***	0.077***	0.198***	0.166***	0.074***
	(0.027)	(0.028)	(0.020)	(0.006)	(0.006)	(0.005)
Panel B: Plead Guilty						
Detained	0.176***	0.162***	0.058***	0.199***	0.165***	0.071***
	(0.027)	(0.028)	(0.020)	(0.006)	(0.006)	(0.005)
Panel C: Minimum Sentence Length (Days)						
Detained	40***	38***	35***	28***	26***	17***
	(3)	(3)	(4)	(1)	(1)	(1)
Time and Courtroom FE	X	X	X	X	X	X
Demographics		X	X		X	X
Criminal History			X			X
F Statistic	2,848	2,987	3,215	1,009	1,080	1,388
Number of Instruments	1	1	1	23	23	23
Observations	726,172	726,167	726,167	725,480	725,480	725,480

Notes: Table 2.7 reports estimates of the effect of pretrial detention on case outcomes for misdemeanors using instrumental variables. The first three columns use the judge-level variation in the propensity to detain pretrial for a given crime type as an instrument for being detained with varying controls. The last three columns use the set of Lasso selected instruments from the full set of constructed candidate instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

unfavorably for misdemeanor cases.

2.5.4 Effects on Charge Reduction

Not only are detainees more likely to plead guilty, but the plea deals they accept are less favorable. Using information on both crimes with which an individual is charged at arraignment (arraignment charges) and the crimes of which he is ultimately convicted (disposition charges), we determine whether the most serious disposition charge belongs to a lower class than the most serious arraignment charge. Crimes are organized into ordered classes based on seriousness. The crimes in our sample fall into eight different classes (class A-E felonies and class A, class B, or classless misdemeanors). Table 2.8 reports estimates of the effect of detention on charge reduction conditional on being convicted. We find evidence of large negative effects on charge reduction, meaning detainees are less likely to be convicted of less serious crimes than the one with which they were charged at arraignment. This effect could translate into more severe punishments for individuals who are subsequently arrested and charged again, because criminal history is an important dimension of sentencing guidelines. Many people who interact with the criminal justice system do so multiple times during the course of their lives. In our sample, only 59% of defendants are first-time offenders, with defendants averaging 3.5 prior felony and misdemeanor convictions. As individuals repeatedly face criminal charges, the negative consequences of being detained even once will continue to accrue.

2.5.5 Effects by Race and Ethnicity

In Tables 2.9 and 2.10, we test for differences in effects across racial and ethnic groups using the IV specification with the relaxed monotonicity assumption. Overall, effects for blacks, whites, and Hispanics are relatively comparable. In both the felony and misdemeanor subsamples, whites have the smallest estimated effects on both conviction and pleading guilty, and Hispanics have the largest effects. However, whites have the largest estimated effects on sentence length. Although we don't see consistent evidence that effects are larger for minorities, blacks and Hispanics are overrepresented in the population of defendants in criminal cases. At the time of the 2010 Census, blacks and Hispanics made up about 23% and 29%, respectively, of the population of New York City in 2010, but 49% and 33% of our sample. Even conditional on being charged with a crime, minorities are overrepresented among detainees (see Appendix Figures B.2 and A.3). As a result, they bear a relatively larger share of the impact of pretrial detention on case outcomes.

Table 2.8: Effect of Detention on Charge Class Reduction

	Felonies			Misdemeanors		
Predicted Incarceration Length	OLS	IV	IVNM	OLS	IV	IVNM
Detained	-0.090***	-0.143***	-0.113***	-0.214***	-0.241***	-0.206***
	(0.003)	(0.022)	(0.006)	(0.002)	(0.017)	(0.006)
F Statistic	-	1,438	848	-	2,824	1,130
Number of Instruments	-	1	27	-	1	23
Observations	167,906	166,467	165,652	486,600	484,212	483,717

Notes: Table 2.8 reports estimates of the effect of pretrial detention on a binary variable equal to 1 if the disposition charge level is lower than the arraignment charge level if convicted (e.g., moving from a class B felony to a class D felony). Columns 1 and 4 use the same specification as columns 4 and 8 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Columns 2 and 5 use the same specification as column 3 of Tables 2.6 and 2.7 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Columns 3 and 6 use the same specification as column 6 of Tables 2.6 and 2.7 and report estimates from an IV estimation that uses the set of Lasso-selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table 2.9: Effects of Detention by Race and Gender for Felonies

	Race/Ethnicity			Gender	
Variables	Black	White	Hispanic	Male	Female
Panel A: Conviction					
Detained	0.130***	0.127***	0.132***	0.130***	0.132***
	(0.008)	(0.017)	(0.009)	(0.006)	(0.016)
Panel B: Plead Guilty					
Detained	0.109***	0.061***	0.117***	0.099***	0.123***
	(0.008)	(0.019)	(0.010)	(0.006)	(0.016)
Panel C: Minimum Sentence Length (Days)					
Detained	144***	181***	155***	161***	131***
	(11)	(22)	(13)	(8)	(13)
F Statistic	675	116	439	1,054	170
Number of Instruments	27	27	27	27	27
Observations	121,977	25,693	83,091	203,937	37,656

Notes: Table 2.9 reports estimates of the effect of pretrial detention on case outcomes for felonies by race and gender. The first three columns report estimates for blacks, whites, and Hispanics. The last two columns report estimates for males and females. The specification used for all columns is the same as column 3 of Table 2.6 and is an IV estimation that uses the set of Lasso selected instruments. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table 2.10: Effects of Detention by Race and Gender for Misdemeanor

	Race/Ethnicity			Gender	
Variables	Black	White	Hispanic	Male	Female
Panel A: Conviction					
Detained	0.070***	0.058***	0.088***	0.077***	0.090***
	(0.007)	(0.018)	(0.009)	(0.005)	(0.013)
Panel B: Plead Guilty					
Detained	0.068***	0.055***	0.084***	0.074***	0.093***
	(0.007)	(0.019)	(0.010)	(0.006)	(0.014)
Panel C: Minimum Sentence Length (Days)					
Detained	13***	20***	23***	17***	14***
	(2)	(4)	(2)	(1)	(4)
F Statistic	876	161	483	1,178	257
Number of Instruments	23	23	23	23	23
Observations	345,422	95,648	240,220	595,059	128,693

Notes: Table 2.10 reports estimates of the effect of pretrial detention on case outcomes for misdemeanors by race and gender. The first three columns report estimates for blacks, whites, and Hispanics. The last two columns report estimates for males and females. The specification used for all columns is the same as column 3 of Table 2.7 and is an IV estimation that uses the set of Lasso selected instruments. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

2.5.6 Possible Mechanisms

In all specifications, the estimated effect of detention on pleading guilty is similar to the effect on conviction. This finding supports the hypothesis that detention influences case outcomes primarily by causing detainees to accept plea bargains more often. Detainees might be more likely to plead guilty for a variety of reasons.

First, some defendants may be offered plea deals that would allow them to go home sooner, potentially without serving any additional time. This is widely accepted among people who work in and write about pretrial detention as a central channel through which detention induces guilty pleas. To investigate whether defendants are pleading guilty to get home sooner, we estimate effects by predicted incarceration length if convicted. To predict incarceration lengths, we first restrict the sample to cases that ended in conviction. Then we regress sentence length on the vector of observable characteristics described in the methodology section. Using the coefficients from this regression, we predict sentence length conditional on conviction for the entire sample. Appendix Tables B.7 and B.8 present estimated effects of detention on conviction by predicted incarceration length. The coefficients on the group with an expected sentence length of zero days are no larger than for groups with positive predicted incarceration lengths.

Thus, although eagerness to get home sooner is a compelling story, it cannot fully explain our findings. Whether we analyze effects by predicted sentence length or by most serious offense charge (Appendix Tables B.9 and B.10), we find strong effects for individuals who would almost certainly face additional time in jail or prison as part of any plea deal. If an individual is not presented with a plea bargain that allows him to go home, why would being detained increase his incentive to accept it? One relevant feature of the criminal justice system is that detainees who are ultimately convicted and sentenced to serve time have the time they spent awaiting adjudication counted against their sentences. This policy lowers the cost of pleading guilty for detainees relative to released defendants, because detainees have paid part of the price of conviction already.

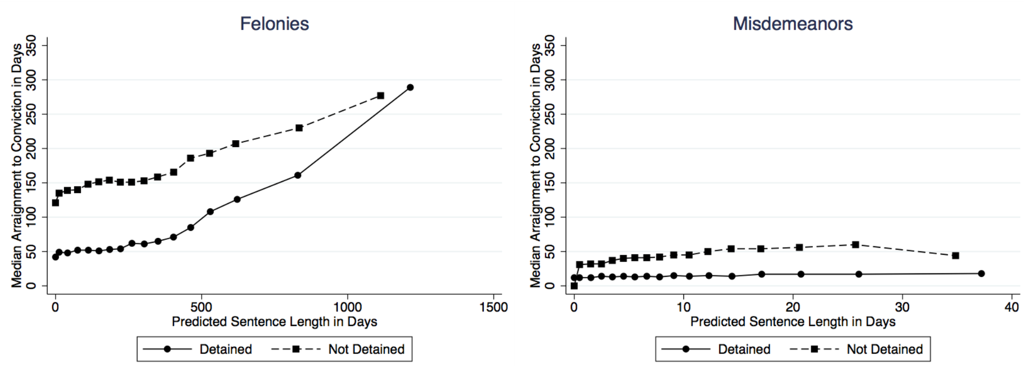
Another possible motivation could be to get moved to a different facility. Pretrial detainees are kept in jail, along with convicts serving relatively short sentences (generally no longer than one year). Prisons are reserved for convicts serving longer sentences, and are designed with long-term residence in mind. Some jails have no yard, no inmate employment, and more limited visiting opportunities. In New York City, the majority of detainees are kept in jails located on Rikers Island, the site of 10 separate jails with a combined capacity of up to 15,000[11]. Rikers has gained notoriety in recent years for abuse and neglect of prisoners. If defendants perceive their detention facility to be worse than wherever they might serve out their sentences, they might opt to plead guilty rather than stay where they are longer than necessary.

This option only increases in its relative appeal if detainees whose cases go to trial are disadvantaged once they get there. Collecting evidence or recruiting witnesses to support one's defense might be more challenging from behind bars. Relatively more detainees end up taking their cases to trial, but the fraction of cases that go to trial is tiny for both groups: only about 1% of cases in the entire sample were adjudicated at trial. By contrast, 34% of released defendants' cases are dismissed, compared to 19% of detainees. These statististics suggest that being detained affects outcomes by causing people whose cases would ultimately have been dismissed to plead guilty, rather than through causing people to plead guilty who would have gone on to be acquitted

[11] Brooklyn, the Bronx, Queens, and Manhattan each have a smaller borough jail, but all together these facilities can only hold up to 3,000 inmates.

at trial if they had been released. It does not follow that expectations about trial outcomes are irrelevant to strategic decision making by defendants (or their attorneys). The criminal justice system moves slowly and unpredictably: for detainees whose cases go to trial, the median time between arraignment and sentencing is 513 days for felonies and 138 days for misdemeanors, with the middle 80% ranging from 226 to 971 days for felonies and 45 to 428 days for misdemeanors. Dismissals often take months: conditional on being dismissed, the median time to dismissal is 188 days for felony cases and 196 days for misdemeanor cases. By comparison, for detainees who plead guilty, the median time between arraignment and sentencing is 80 days for felonies and 15 days for misdemeanors. Over time, the prospect of attaching an end date to the period of incarceration may become very attractive in contrast to the alternative of waiting an indeterminate amount of time behind bars for the process to play itself out, risking a trial at which the consequence of losing is a more severe punishment than any plea deal.

Figure 2.5: Time to Conviction for Plead Guilty Cases by Detainment Status



Notes: Each graph shows the median number of days from arraignment to conviction for cases in which the defendant pleads guilty. Each points represents the median for the corresponding bin.

Figure 2.5 shows the relationship between predicted sentence length and the average amount of time between arraignment and conviction for defendants who ended up pleading guilty. One thing that is immediately clear from these graphs is that detainees plead guilty much earlier than defendants who are released pretrial. Detainees accused of felonies wait more than twice as long as

Table 2.11: Effects by Predicted Incarceration Length and Gender for Misdemeanors

	Men			Women		
Predicted Incarceration Length	OLS	IV	IVNM	OLS	IV	IVNM
0 Days	0.016**	-0.483*	0.023	0.118***	-0.819*	0.108*
	(0.008)	(0.253)	(0.033)	(0.014)	(0.479)	(0.058)
1-11 Days	0.103***	0.173**	0.053***	0.148***	0.279**	0.058
	(0.004)	(0.078)	(0.017)	(0.010)	(0.126)	(0.037)
11+ Days	0.104***	0.083***	0.084***	0.106***	0.150***	0.051***
	(0.002)	(0.019)	(0.006)	(0.007)	(0.048)	(0.016)

Notes: Table 2.11 reports estimates of the effect of pretrial detention on conviction for misdemeanors by the predicted length of incarceration if convicted and by gender. Columns 1 and 4 use the same specification as column 8 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Columns 2 and 5 use the same specification as column 3 of Table 2.7 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Columns 3 and 6 use the same specification as column 6 of Table 2.7 and reports estimates from an IV estimation that uses the set of Lasso selected instruments. Corresponding estimates for felonies are shown in Table A.11. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

those charged with misdemeanors before pleading guilty, which is consistent with a more gradual process of becoming discouraged. For the misdemeanor subsample, the median time between arraignment and disposition is less than three weeks at every predicted sentence length. By contrast, the median gap between arraignment and disposition is almost 50 days for felony defendants in the lowest predicted sentence-length bin, and grows larger for groups with longer predicted sentence lengths.

Pretrial detention might affect some defendants more adversely than others. For individuals with dependents, detention could also upend their children's lives, possibly putting them into the foster system. This scenario could create a powerful incentive for defendants to avoid detention. We find suggestive evidence supporting this possibility. Because felonies are, by definition, crimes that are punishable by at least a year in prison, we would not expect defendants charged with a felony to have the option of pleading guilty and returning home right away. However, many misdemeanor defendants would have access to a plea deal that requires little to no additional time in jail. Women are also more likely to be single parents than men: about three quarters of all single-parent households are headed by women (Livingston, 2013). If defendants with dependents are even more likely to plead guilty if doing so secures their immediate release, we would expect to see a larger disparity between coefficients for men and women with low expected incarceration lengths. In Table 2.11 we analyze effects for the misdemeanor sample by gender and predicted sentence length. For defendants with a predicted sentence length of zero days, effects are noticeably stronger for women. The difference disappears for individuals with longer predicted sentences. Although not conclusive, these estimates are consistent with the hypothesis that low-income defendants with children are especially disadvantaged by the bail system.

2.6 Conclusion

Taken together, our results indicate a strong causal relationship between pretrial detention and case outcomes. Across specifications and subgroups, we see consistent evidence that detainees plead guilty more often to more serious offenses, and some evidence that they serve longer sentences. The financially disadvantaged bear the brunt of these effects, because the majority of detainees are held after failing to post bail at relatively low levels. Our findings suggest several policy avenues for improving the criminal justice system. Getting rid of money bail entirely would eliminate the unfair disadvantage to individuals with limited financial resources. This reform could, in theory, be implemented without changing the fraction of defendants that are detained pretrial, but policymakers should seriously consider whether the costs of the current detention rate outweigh the benefits. The main legal motivations for pretrial detention are to ensure appearance at court and public safety. The optimal detention rate is unlikely to be zero, but may be much lower than the status quo. Some parts of the country have already implemented bail-reform measures. For example, D.C. has completely eliminated money bail in favor of an "in or out" court system and enhanced pretrial services. Only about 15% of accused persons are detained. Of those released, about 12% are re-arrested pretrial, but fewer than 1% of these persons are alleged to have committed a violent crime. About 88% return to court (Keenan, 2013). Avoiding costs associated with re-arrests and failures to appear is an important benefit of pretrial detention, particularly for high-risk individuals, but our analysis suggests policymakers will severely underestimate the costs of detention if they do not figure in the effect on case outcomes.

Although we explore the mechanism behind the effects, additional work is needed to more accurately pinpoint the relative importance of the different channels. Another open question is how the effects of pretrial detention on case outcomes vary as the marginal defendant changes. Our results are local effects based on current detention practices in New York City. If the criminal justice system were to relax or tighten its criteria for pretrial detention, estimated coefficients would pick up effects for a different set of marginal defendants. Understanding how effects vary across the distribution of defendants would be an important component in determining optimal detention practices.

CHAPTER 3
THE HABIT-FORMING EFFECT OF SUBSIDIES: EVIDENCE FROM WIC[1]

3.1 Introduction

Public assistance to poor households through cash transfers allows recipients to select their own utility maximizing bundles without excessive constraints and provides relatively low administrative costs. However, the welfare system in the United States includes a broad set of vouchers that can only be used for specific goods and services. Vouchers that apply to a broad class of items, such as through the Supplemental Nutrition Assistance Program (SNAP), are effectively equivalent to a cash transfer for most households (Hoynes and Schanzenbach, 2009). In contrast, the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) provides recipients a set of vouchers for specific food items. In this paper, we examine whether programs that provide highly targeted vouchers for a sustained period continue to influence behavior after the program ends. The process of consumption patterns in one period positively influencing consumption patterns in future periods is a form of habit formation (Rabin, 2013). Habit formation can occur through at least two mechanisms, both of which are possible with WIC vouchers. First, 'classic' habit formation occurs when individuals become used to eating particular foods, and this becomes an automatic pattern of behavior. Second, consuming certain foods may influence an individual's food preferences through repeated exposure (Birch and Marlin, 1982).

Several recent studies provide evidence that short-run programs that influence food choices in schools can continue to affect behavior once the program ends through habit formation (Belot

---

[1]This chapter is coauthored with David Frisvold and Joseph Price. The authors are grateful for research assistance provided by Michael Gmeiner, Chelsea Hunter, Nathalia Myrrha, Adam Shumway, and Jacob Walley. The authors thank Laura Argys, Kevin Gee, Adriana Lleras-Muney, David Johnson, Daniel Tannenbaum, Nathan Tefft, seminar participants at Georgia State University, UC Davis, UC Merced, and University of Illinois, and participants at the American Society of Health Economists, Association for Public Policy Analysis and Management, Southern Economic Association, Western Social Science Association, and Western Economic Association conferences for helpful comments. Funding for this project was made possible in part by grant number 1H79AE000100-1 to the UC Davis Center for Poverty Research from the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Analysis (ASPE), which was awarded by the Substance Abuse and Mental Health Services Administration (SAMHSA). The views expressed are those of the authors and do not necessarily reflect the official policies of the Department of Health and Human Services. This paper uses data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen.

et al., 2013; List and Samek, 2015; Loewenstein et al., 2016).[2] Our paper extends this work in several ways. First, we analyze the effects of several years of exposure to incentives, while previous studies measure effects for programs that last only a few weeks. Second, we test for habit formation throughout a longer follow-up period than most existing studies. Third, we test for habit formation in a context where parents are the decision makers, rather than children. Finally, we examine habit formation as a result of a national public program, in contrast to a local field experiment. Thus, we also contribute to the literature on food assistance programs and antipoverty programs more generally by examining whether targeted public vouchers continue to influence behavior after eligibility ends.

To test for habit formation through WIC vouchers, we use detailed scanner data from the Nielsen Consumer Panel. These data allow us to examine changes in household purchases of WIC-eligible and WIC-ineligible products that occur during program participation and whether these changes persist after children in the household age out of eligibility. To identify the effects of WIC vouchers, we take advantage of the changes in the specific items covered by WIC that occurred in 2009, particularly the introduction of whole grains and produce to the package of goods offered to children. We begin by verifying that the package revision affected what households purchased while the youngest member of the household was eligible for WIC. We find that income-eligible households with at least one age-eligible child increase purchases of whole grain products by 15 percent after the WIC package change. We also find that the package change does not significantly influence produce purchases, which is perhaps due to the size of the voucher relative to the amount of produce that WIC-eligible households were purchasing prior to the package change. Increasing the confidence that the results are due to the WIC package change, we do not find corresponding changes in whole grains or produce purchases at the time of the package change among households in which the youngest child is older than the age threshold for WIC eligibility.

Next, we examine whether aging out of WIC eligibility influences household purchases. We find that income-eligible households decrease whole grain purchases after the youngest member of the household turns 5 years old and is no longer eligible to receive WIC. Within six months of aging out of the household, the magnitude of the decrease in whole grain purchases is similar to the magnitude of the increase in whole grain purchases from the introduction of whole grain products

[2]Another branch of the literature studies persistence in response to exercise programs, finding evidence that short-term financial incentives continue to influence behavior for several weeks after incentives are removed (Charness and Gneezy, 2009; Royer et al., 2015).

to the WIC package.

The sharp change in the WIC package creates a natural experiment in which some households aged out of WIC eligibility just before the package change while other households were exposed to the new WIC items for varying amounts of time based on the age of their youngest child when the package change occurred. We use the variation in the length of program eligibility after the package changes to determine the persistent impact of the vouchers on household purchases. This provides a test of whether the length of time that a household is influenced to adopt a healthy purchasing pattern can create a long-run change in household purchases. We do not find that households exposed to whole grain and produce vouchers for longer are more likely to purchase these products after the youngest member of the household ages out of eligibility.

Overall, the results indicate that the WIC package revision increased purchases of whole grain products, which was one of the goals of the repackaging. However, we find little evidence that these changes in purchases of WIC items continue for the long-term after households are no longer eligible for WIC.

3.2 Background

The objective of WIC is to supplement nutrient intake for pregnant and postpartum women and young children. WIC is a federal program that was permanently authorized in 1974 and operates through state and local agencies. In 2014, 8.3 million people received WIC benefits every month; more than half of the recipients were children aged two to four years (USDA, 015a). The total costs of the program were $6.3 billion and the average monthly benefit per participant was $43.64 in 2014 (USDA, 015b). Eligibility for WIC is determined at the individual level, primarily based on age and household income. Women may be eligible if they are pregnant, up to six months postpartum, or 7 to 12 months postpartum and breastfeeding. Children may be eligible up to their fifth birthday. Individuals are eligible if their household income is below 185% of the poverty line, or if they are participating in the federal aid programs Temporary Assistance to Needy Families (TANF), the Supplemental Nutrition Assistance Program (SNAP), or Medicaid.[3] Since we are using data on household food purchases, we refer to households as WIC eligible if they contain at least one eligible

[3]Children must also be deemed at nutritional risk to be eligible. WIC's definition of nutritional risk includes conditions such as being overweight, underweight, or anemic. It can also include the characteristics of the mother including age, inadequate diet, and past pregnancy complications. Bitler et al. (2003) find that, in practice, the nutritional risk criteria does not prevent otherwise eligible individuals from receiving WIC benefits.

Table 3.1: Revisions to the WIC Package for Children Ages 2 through 4

	Maximum Allowance Before 2009	Maximum Allowance After 2009
Whole Grains	-	32 oz.
Produce Voucher	-	$6
Any milk	24 qt.	-
Reduced-fat milk	-	16 qt.
Cheese	64 oz.	16 oz.
Juice	288 oz.	128 oz.
Cereal	36 oz.	36 oz.
Eggs	30	12
Dried beans	16 oz.	16 oz.
Canned beans	-	64 oz.
Peanut butter	18 oz.	18 oz.

Notes: The WIC package for one-year-old children is the same as the package described in this table, except for the milk allowance. The produce voucher was increased to $8/month in 2014. Before 2009, up to 4 lbs. of cheese could be substituted for milk at a rate of 1 lb. cheese per 3 quarts milk. After 2009, up to 1 lb. of domestic cheese with reduced sodium, fat, or cholesterol may be substituted for milk at a rate of 1 lb. cheese per 3 quarts milk. After 2009, half of the cereals on each state's authorized list must have whole grain as the primary ingredient by weight). After 2009, canned beans are allowed as a substitute for dried beans. Peanut butter is a substitute for dried beans. Sources: USDA (2011, 2016a,b)

member.

The package of foods subsidized for WIC participants can differ for women who are pregnant, postpartum, and breastfeeding, infants who are younger than 12 months, and children who are younger than 60 months. These packages remained largely unchanged between 1972 and 2009. The review of the WIC package began in 2003, and the packages were revised in 2007 to be consistent with the Dietary Guidelines for Americans. All states were required to implement the revised packages by October 1, 2009. Most states implemented these revisions on October 1. A few states implemented the revisions earlier in the summer, and New York and Delaware implemented the revisions in January.[4] Table 3.1 provides a comparison of the original and revised food packages for children. The revisions included the addition of whole-grain products and produce to the package. The types of milk included in the package changed to require low- or non-fat for children two and older, and the quantity was reduced. Similarly, for cheese, which is a dairy substitute for milk, the types allowed changed and the quantity was reduced. The types of cereal changed to emphasize products with whole grains, while the quantity included remained the same. The quantities of juice and eggs included in the package were reduced. Our analysis focuses on the specific products and quantities that are included in the WIC packages and the changes in the included products that occurred in 2009. We focus primarily on whole grains and produce, because these products were new to the WIC package in 2009, but show the results for all products in the appendix.

A number of papers have examined the influence of the 2009 WIC food package revisions on the availability of specific food items in stores, prices, purchases, and consumption. These studies compare characteristics of the supply side (availability and prices) before and after the package change or the behavior of WIC participants (purchases and consumption) before and after the package change. The results from these studies consistently show that the availability of healthier foods included in the WIC package increased in stores (Andreyeva and Luedicke, 2013) with some evidence that the prices of these foods did not rise or decreased (Zenk et al., 2014). Further, these studies consistently find that the WIC package revision increased purchases and consumption of the foods included in the new package, resulting in an overall improvement in diet (Whaley et al., 2012; Andreyeva and Luedicke, 2013).

Our paper makes two significant contributions to the literature on the influence of WIC participation on food purchases. First, we use a more credible research design to determine the impact of the WIC program by comparing purchasing patterns before and after the WIC package change in 2009 for eligible and non-eligible households and also exploiting variation in the age of the youngest child across households to incorporate the age eligibility requirements for WIC participation. Second, we examine whether exposure to specific WIC-eligible items has a persistent influence on what WIC participants purchase even when they are no longer eligible for WIC. This contributes to the literature on habit formation by examining a targeted voucher that changes what individuals purchase for several years using data that allows us to continue to follow the purchasing

[4]The implementation dates for states that introduced the revised WIC packages before October 1, 2009 are: January 2, 2009 for Delaware and New York; May 1 for Kentucky; June 1 for Colorado; July 1 for Utah; August 1 for Kansas, Michigan, Minnesota, Oklahoma, Oregon, and Wisconsin; August 3 for Illinois; September 1 for South Dakota; and September 28 for Alabama, Arkansas, and Indiana.

patterns of the household months after leaving the program.

3.3 Data

The analysis in this paper is based on Nielsen Homescan Consumer Panel data for the years 2004 through 2014, which include detailed data on food items purchased for about 40,000 or more households each year. Participants in the panel are given a special scanner that they use to scan in all items purchased at any grocery store along with all of the information recorded on the store receipt. The scanner records the UPC code for each item. At the end of each week, the household transmits their data to Nielsen and receives points, which can be exchanged for merchandise in a way similar to using a credit card. This system creates a strong incentive for households to upload their purchase data each week (Harding et al., 2012).

Since data are recorded at the UPC level, they include sufficient product characteristics to determine the type of item being purchased and the exact amount purchased. We use these detailed data to look at specific items that are included as part of the WIC bundle, along with their quantities. We primarily examine whole grains (whole-wheat bread and bread substitutes) and produce (fruits and vegetables), which are the products most affected by the package revisions. In the appendix, we also examine milk, breakfast cereals, cheese, eggs, and fruit juice. Further, we show the results for fish in the appendix as a falsification exercise. Fish is included in the WIC package for fully breastfeeding women, but not for children. All products are measured in ounces per month. In the appendix, we also report results for produce in dollars spent since WIC provides a dollar voucher for produce instead of a quantity voucher.[5]

The Nielsen data also provide socio-economic characteristics of the household, including household income, age of family members, household size, education levels, and other demographic characteristics. Two of the most important characteristics that determine whether a household member is eligible for WIC are household income and the age of the household's youngest child. Eligible households are likely to have income less than 185% of the federal poverty guidelines.[6]

[5]Our food measures include all products in each category and not just the specific brands or sizes that are included in vouchers for each state. Whole grain products include whole-wheat and whole-grain bread, buns, rolls, and noodles; corn and wheat tortillas; brown rice; bulgur; oats and oatmeal; and barley. Produce includes all canned, frozen, and fresh fruits and vegetables and dried fruits. Low-fat milk includes skim and up to 2 percent non-sweetened milk. Whole milk includes non-sweetened whole and evaporated milk. Cereal includes hot and cold cereals with separate categories for whole grain cereals and sugar cereals. Cheese includes all imported and domestic cheeses and string cheeses, but not spreads, cream cheese, or cheese dips. All eggs are included. Juice includes fresh and frozen juice and juice drinks of any size.

[6]States have the flexibility to set the income eligibility threshold between 100 and 185 percent of the

Until 2011, the Nielsen data reported total household income in the full calendar year that is two years prior to the data on purchases. In the fall preceding the panel year, households were asked to report their total annual income for the previous year. In 2011, Nielsen changed the question about income to ask households to report their estimated annual income at the time of the survey. Since Nielsen believed that households were reporting their current estimated annual income instead of referring to the prior year's tax returns, this change should have increased the consistency of the reporting period of income (Kilts Center for Marketing, 2014). Household income is measured as a categorical variable that has rather narrow bands at the lower income levels.[7] Using the minimum of these income bands and the number of household members, we are able to approximate where a household stands in relation to the poverty guidelines. In the sample of households with reported income, about 20% of the households fall below 185% of the federal poverty guidelines in any given year.

Households are only eligible for WIC if they have a child under the age of five (or if the mother is pregnant). We use the birth month and year of household members to determine age eligibility for WIC. There is no information about whether women in the household are pregnant. As a result, we use information from subsequent years of the household to infer the timing of pregnancy. Some households with infants born in their final year in the panel may be incorrectly marked as ineligible for WIC when they are eligible. However, this will just bias against finding an impact of being on WIC during the WIC-period but not influence our estimates of how long the family is on WIC when we examine post-WIC behavior.

Beginning in 2006, the Nielsen data include variables describing whether the household is currently receiving and has ever received WIC. Kreider et al. (2016) and Bitler et al. (2003) document that WIC participation is generally underreported in survey data. The low levels of self-reported WIC participation in the Nielsen data are consistent with rampant underreporting. For example, WIC serves over half of all infants born in the United States (citation), but less than 10% of Nielsen households with infants report receiving WIC. Because missing responses are indistinguishable from

federal poverty guidelines, but all states use the maximum amount of 185 percent. Also, pregnant and post-partum women, infants, and children under 60 months are categorically eligible if they also participate in SNAP, Medicaid, or TANF. However, only about two percent of WIC participants report household income above 185 percent (GAO, 2013).

[7]The first few income categories are: under $5,000, 5,000-7,999, 8,000-9,999, 10,000-11,999, and 12,000-14,999. The income categories are $5,000 apart from $20,000 through $50,000 and then $10,000 apart through $70,000.

negative responses for WIC participation in the Nielsen data, self-reported participation is likely to be even less reliable than in other surveys. Due to the low quality of the self-reported WIC information, we focus on eligibility rather than participation.

We treat a household as WIC eligible if household income is below 185% of the federal poverty guidelines and the youngest member of the household is less than 60 months old. Our analysis sample excludes households with income above 400 percent of the federal poverty guidelines in all waves of the panel and households that did not have a child under 60 months of age for at least one month of the panel. These restrictions yield 9,784 households with an average duration in the panel of 46.5 months for a total of 455,772 household-months in our analysis sample.

Table 3.2 provides descriptive statistics for all households, WIC-eligible households, and WIC-ineligible households. There are 93,713 observations in which the household is eligible for WIC, based on the age of the youngest member of the household being less than 60 months of age and income being less than or equal to 185 percent of the poverty guidelines. Average household income is $30,130 and the average household size is 4.89 persons. There are 372,422 household-month observations of households with income greater than 185 percent of the poverty guidelines or without an age-eligible child, which are ineligible for WIC during that month. Since the analysis sample is restricted to households with at least one age-eligible child for at least one month, these observations include households who formerly received WIC benefits. For these observations, average household income is $64,960 and the average household size is 3.89 persons.

WIC-eligible households purchase similar quantities of whole grains as WIC-ineligible households. The revised WIC package includes vouchers for 32 ounces of whole grains per eligible child. Although not shown in the table, prior to the package revision, WIC-eligible households purchased 23.5 ounces per month (with a standard deviation of 48.6). After the revision, WIC-eligible households purchased 30.0 ounces per month (with a standard deviation of 54.4).

Although they spend $19.83 per month on average on produce, which is slightly less than ineligible households, WIC-eligible households purchase 190.1 ounces of produce per month compared to 181 ounces for ineligible households. The produce voucher in the revised WIC package is $6 per month per eligible child, which is significantly less than the average WIC-eligible household spends on produce. The average monthly expenditures for WIC-eligible households was $16.86 (with a standard deviation of $16.65) before the package change and is $23.92 (with a standard deviation of $24.49) after the package change. The produce voucher targets expenditure amounts, not quantities, and expenditures on produce were higher after the package change. In contrast, the average ounces

Table 3.2: Summary Statistics

Food purchases (oz.):	All		WIC-Eligible		WIC-Ineligible	
Whole Grains	26.08	(47.79)	25.59	(49.92)	26.19	(47.29)
Low-Fat Milk	382.10	(474.39)	374.84	(493.17)	383.73	(470.07)
Whole Milk	121.60	(266.22)	155.89	(309.54)	113.92	(254.89)
Total Milk	503.70	(520.09)	530.73	(558.83)	497.65	(510.82)
Produce	182.68	(195.10)	190.12	(209.66)	181.02	(191.65)
Cereal	58.08	(75.22)	61.20	(76.77)	57.38	(74.85)
Whole-Grain Cereal	15.62	(33.48)	16.47	(32.07)	15.43	(33.79)
Sugar Cereal	29.51	(47.08)	31.52	(49.96)	29.06	(46.40)
Cheese	47.07	(53.22)	49.48	(56.82)	46.53	(52.37)
Eggs	47.51	(64.44)	50.03	(67.10)	46.94	(63.81)
Juice	296.76	(377.02)	301.36	(386.98)	295.73	(374.74)
Expenditures ($):						
Food Expenditures	314.60	(229.03)	321.41	(249.15)	313.07	(224.26)
Produce Expenditures	21.43	(22.31)	19.83	(20.77)	21.78	(22.63)
Characteristics:						
Household Income ($1000)	58.59	(30.22)	30.13	(12.87)	64.96	(29.31)
Household Size	4.07	(1.52)	4.89	(1.59)	3.89	(1.44)
Age of the Youngest (Yrs.)	10.16	(15.96)	2.18	(1.58)	11.95	(17.14)
White	0.72	(0.45)	0.72	(0.45)	0.72	(0.45)
Black	0.11	(0.32)	0.13	(0.33)	0.11	(0.31)
Hispanic	0.10	(0.30)	0.10	(0.30)	0.10	(0.29)
Other Race/Ethnicity	0.07	(0.25)	0.06	(0.23)	0.07	(0.25)
Married	0.81	(0.39)	0.75	(0.43)	0.83	(0.38)
High School or Less	0.15	(0.36)	0.25	(0.43)	0.13	(0.33)
Some College	0.28	(0.45)	0.35	(0.48)	0.27	(0.44)
College Graduate	0.57	(0.50)	0.40	(0.49)	0.61	(0.49)
N	455,772		83,350		372,422	

Notes: The unit of observation is household-month. Standard deviations appear in parentheses to the right of the mean values. The sample includes 9,784 unique households. The race/ethnicity and marital status variables refer to the household head. The education variables reflect the highest degree of schooling of an adult in the household.

purchased by WIC-eligible households fell from 193.5 to 189.0 ounces per month.

WIC-eligible households purchase approximately similar quantities of low-fat milk, cereal, cheese, eggs, and juice as WIC-ineligible households and purchase higher quantities of whole milk. Average food expenditures are $321.41 per month for WIC-eligible households, which includes purchases using WIC voucher, and are $313.07 per month for WIC-ineligible households.

3.4 Analysis

The objective of our analysis is to estimate the extent to which WIC affects household purchasing patterns and whether this change persists after eligibility ends. Comparisons of the purchasing behavior of households with and without individuals eligible for WIC are unlikely to yield consistent estimates, since the unobserved characteristics that are related to participation are also likely to affect purchasing decisions. Thus, our analysis consists of a series of difference-in-differences regressions that make use of the 2009 package revision and the age-eligibility criteria. First, we estimate whether changes in the items included in the WIC packages in 2009 affected purchasing patterns for WIC-eligible households. Second, we examine changes in purchasing patterns before and after the package change when households lose eligibility because the youngest member of the household turns 5 years old. Third, we estimate whether the amount of time that WIC-eligible households are exposed to the specific items in the new package vouchers affects their purchases of these items after these households are no longer eligible for WIC, based on the age of the youngest child.

3.4.1 Changes Due to the Revision of the WIC Packages

We begin by estimating changes in household purchasing patterns that occurred in 2009 when the items included in the WIC package changed, using the implementation date for the state where the household was living. We compare the changes in the amount purchased of specific product categories before and after the package changes for WIC-eligible households to changes that occurred in households that were ineligible for WIC. Specifically, we estimate the following difference-in-differences specification:

$$Y_{ht} = \beta_0 + \beta_1 Inc_{ht} + \beta_2 After_t + \beta_3 Inc_{ht} * After_t + \gamma X_{ht} + \rho_h + \delta_t + \epsilon_{ht} \tag{3.1}$$

where Y_{ht} denotes the amount purchased of a specific product for household h in month t, between January 2004 and December 2013. Inc denotes whether the household is income-eligible for WIC.

We initially restrict the sample to households with at least one age-eligible child. As a result, Inc is equivalent to WIC eligibility. $After$ is a dummy variable for whether the purchase occurs after the package change based on the date of the package changes in the state of residence of the household. ρ represents household fixed effects. δ represents time (month and year) fixed effects, which control for any annual trends in purchasing patterns, seasonality of purchasing patterns within the year, and any changes in the reporting patterns of household purchases in the survey. X is a vector of household characteristics including age of the youngest person in the household, income, household size, race/ethnicity, marital status, and educational attainment. β_3 is the parameter of interest; it measures the change in purchases after the package revision for WIC-eligible households compared to income-ineligible households with an age-eligible child.

Table 3.3 displays the estimates of equation (1) for whole grains and produce purchases, which are the two categories added under the new guidelines.[8] The top panel displays results for households with at least one age-eligible child who is at least 12 months old, which includes 9,342 households and 180,796 household-month observations.[9] We find that, after the package change, WIC-eligible households increased purchases of whole-wheat products by 3.5 ounces, which is a 14.8% increase relative to the mean of 23.5 ounces of income-eligible households prior to the package change. This suggests that the addition of the whole-grain category to the WIC package had a significant positive effect on purchases of whole grain items.[10]

The coefficients for produce purchased and the amount of money spent on food are both positive, but imprecisely estimated and small in magnitude. The coefficient on produce represents a

[8]The results for additional outcomes are shown in Appendix Table C.1. The package revision did not change the maximum allowances or product types for these products as significantly as the products featured in Table 3.3. The estimates are mostly small and statistically insignificant for WIC-eligible households, with the exception of juice, which increased by 11.8 ounces or 3.7 percent. The revised package decreased the amount of juice and changed the allowed types to exclude juice drinks. In results not shown, purchases of juice products included in the revised WIC package decreased after the package revision, consistent with the smaller allowance in the package, but purchases of juice drinks increased after the package revision. Fish is included as a falsification exercise since fish is included in the revised package for post-partum breastfeeding mothers, but not children.

[9]The results are similar when we restrict the sample further to include the 4,466 households that are in the sample for at least one month before and after the WIC package change, which includes 108,281 household-month observations.

[10]The results are not sensitive to whether we control for household characteristics. Appendix Table C.2 displays the results from equation (1) that do not control for household characteristics and additional regressions that use each household characteristic as an outcome variable. The point estimates for all outcomes are small in magnitude relative to the sample means for income-eligible households prior to the package change and are not statistically significant, except for parents' marital status.

Table 3.3: Difference-in-Differences Estimates of the Impact of the WIC Package Change on Household Purchases

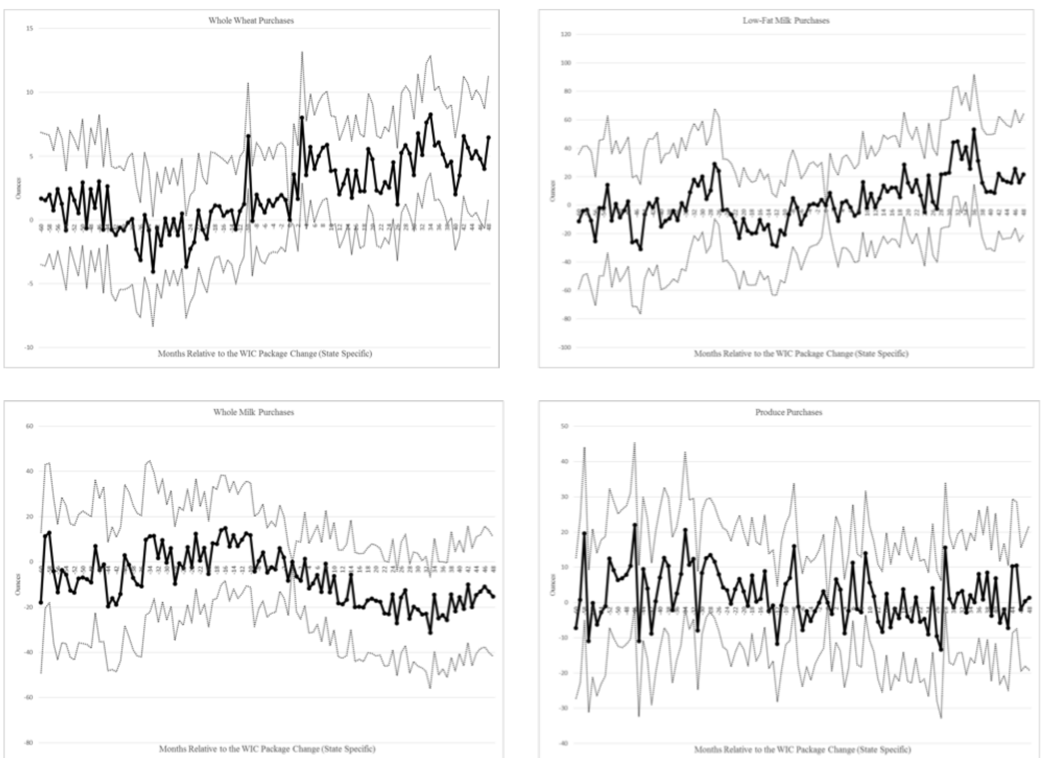
	Whole Grains	Produce	Food ($)
Panel A: Sample is Households With At Least One Age-Eligible Child			
Income-Eligible X After Package Change	3.485	1.035	9.943
	(1.228)	(4.592)	(6.031)
Income-Eligible	-0.394	-4.574	-19.335
	(1.125)	(3.736)	(4.983)
After Package Change	-1.711	-6.827	-11.509
	(0.862)	(3.095)	(3.653)
Pre-Revision Mean	23.48	193.48	302.16
Observations	180,796	180,796	180,796
Panel B: Sample is Households Without Any Age-Eligible Children			
Income-Eligible X After Package Change	1.969	2.108	6.237
	(1.484)	(5.360)	(7.158)
Income-Eligible	-2.181	-0.463	-13.200
	(1.782)	(5.236)	(6.469)
After Package Change	0.490	-0.251	-2.437
	(1.181)	(3.725)	(4.085)
Pre-Revision Mean	22.58	185.50	278.11
Observations	93,035	93,035	93,035

Notes: Each column displays estimates from separate regressions. In Panel A, the sample is restricted to households in which the youngest child is between 12 and 59 months of age. In Panel B, the sample is restricted to households in which the youngest child is at least 60 months of age. Standard errors are in parentheses and clustered within households. Each regression also includes controls for age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school or high school, some college; college graduate omitted), and year, month, and household fixed effects. Whole grains and produce are measured in ounces per month. The pre-revision means are the means prior to the package change for income-eligible households for each sample.

0.5 percent increase relative to the mean prior to the package revision for WIC-eligible households and the coefficient on food expenditures represents a 3.2 percent increase. Since WIC-eligible households already purchased more than $6 of produce each month, the estimates for produce are consistent with the possibility that these households were able to substitute the value of this voucher to purchase other products. Although produce was included in the revised package, the voucher for $6 was well below the pre-revision mean of $16.86 spent on produce, and over 70 percent of household-month observations were above $6 for WIC-eligible households prior to the package revision. In contrast, the quantities of whole grains allowed in the revised package (32 ounces) were above the pre-revision mean for WIC-eligible households (23.5 ounces).

The identifying assumption for equation (1) is that the changes over time in food purchases of income-ineligible households with an age-eligible child would be similar to the changes for WIC-

Figure 3.1: Difference in Purchases for WIC-eligible and WIC-ineligible Households Before and After the Package Change



Notes: These figures show the differences in purchases for whole-grain products, low-fat milk, whole milk, and produce for WIC-eligible households compared to WIC-ineligible households by month relative to the WIC package change for the state of residence.

eligible households in the absence of the WIC package revision. To assess the credibility of this assumption, we examine the monthly food purchasing trends of these two groups of households before and after the package revision. As shown in Figure 3.1, the differences in whole grain purchases for eligible and ineligible households are similar prior to the package change. Also, in the bottom panel of Table 3.3, we estimate the effect of the package change for households without any age-eligible children as a falsification test. These households aged out of WIC eligibility prior to the package revision and would not have been exposed to whole grain or produce vouchers. We do not find statistically significant impacts on whole grains, produce, or food expenditures from the WIC package revision for this sample. Further, the point estimate for whole grains is much smaller for this sample at 1.97 ounces.

Overall, our estimates show that the package revision changed household purchasing patterns for whole grain products, suggesting that the specific products included in the WIC vouchers can be in important policy tool for influencing the types of foods low-income families with children purchase. At the same time, they confirm that providing vouchers that are inframarginal to pre-program spending may be ineffective in changing behavior (beyond the effects associated with a cash transfer). Our results also demonstrate that the WIC package change is a useful context for studying habit formation, since a change in the WIC package is able to successfully change the household purchase patterns while the household is eligible for WIC. Thus the WIC package creates a similar type of behavioral change that accompanies the use of incentives in school food environments. An advantage of using the behavioral change that accompanies the package change is that it creates a much longer period of change than is normally possibly with field experiments in schools.

3.4.2 Persistence of Effects after Losing Age Eligibility

Next, we estimate the changes in household purchasing patterns when income-eligible households lose eligibility after the youngest child reaches 60 months of age. Specifically, we estimate:

$$Y_{ht} = \alpha_0 + \alpha_1 Inc_{ht} + \alpha_2 * 1[Age \geq 60]_{ht} + \alpha_3 Inc_{ht} * 1[Age \geq 60]_{ht} + \theta X_{ht} + \rho_h + \delta_t + \epsilon_{ht} \quad (3.2)$$

where $1[\cdot]$ is an indicator function and $1[\text{Age} \geq 60]$ denotes that the youngest member of the household is not age-eligible for WIC (at least 60 months old). All other parameters and coefficients are defined analogously to those in equation (1). α_3 is the coefficient of interest, which represents the change in purchases upon losing WIC eligibility for income-eligible households compared to

Table 3.4: Difference-in-Differences Estimates of the Impact of Aging out of WIC Eligibility

	Before the	After the Package Change					
	Package Change	All Periods	First 3 Months	First 6 Months	First 9 Months	First 12 Months	First 24 Months
Whole Grains	0.769	-3.954	-2.604	-3.521	-3.992	-3.682	-4.004
	(1.186)	(1.256)	(1.491)	(1.416)	(1.429)	(1.414)	(1.370)
Produce	-8.095	-6.211	-5.953	-6.787	-4.642	-3.603	-2.911
	(4.267)	(4.049)	(5.643)	(4.933)	(4.827)	(4.690)	(4.367)
Food ($)	-6.647	-11.902	-2.818	-5.445	-6.654	-5.708	-7.374
	(4.385)	(6.733)	(6.669)	(6.602)	(6.592)	(7.002)	(7.133)
N	133,882	139,949	91,243	95,972	100,346	104,404	118,838

Notes: Each row and column displays estimates from a separate regression. These estimates are the coefficients corresponding to the interaction of income-eligibility (under 185% FPL) and having the youngest member of the household at least 60 months old (no longer age-eligible). Standard errors are in parentheses and clustered within households. The sample is restricted to households where the youngest child is within 48 months of turning 60 months of age. The first column displays estimates for households in which the youngest member of the household aged out of WIC eligibility prior to the package change. The second column displays estimates for households in which the youngest member of the household aged out of WIC eligibility after the package change. The sample for the third column is a subset of the sample from the second column that is restricted to households in which the youngest member is not older than 63 months of age (within the first 3 months of aging of out WIC eligibility). Each subsequent column has a sample that is constructed similarly. Each regression also includes controls for age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school or high school, some college; college graduate omitted), and year, month, and household fixed effects. Whole grains and produce are measured in ounces per month.

income-ineligible households.

To examine whether there is a persistent impact of WIC vouchers and how long the impact persists, we estimate equation (2) with periods of different lengths after aging out of WIC eligibility. These results are shown in Table 3.4. The first column estimates effects for households that aged out of eligibility prior to the package change. These households that lost eligibility prior to the package change never received any WIC vouchers specifically for whole grains or produce. As a result, any decrease in purchases of these items of losing WIC more generally. The coefficient for whole grains is small and not statistically significant, suggesting that there is little income effect on whole grain purchases prior to the package change. In contrast, there is a reduction in produce purchases of 8.1 ounces as income-eligible households age out of WIC eligibility, which suggests that the income effect associated with losing the implied value of the WIC package reduces produce purchases.

The rest of Table 3.4 restricts attention to households whose youngest child turned five after the package change such that the household had some potential exposure to the revised set of vouchers. In the second column, we define the post-treatment period to include all months within four years after the youngest turns five. We find that after income-eligible households age out of WIC eligibility, their whole grain purchases decrease by about four ounces per month. This completely undoes the effect of the package change during eligibility as estimated in Table 3.3. However, it also obscures the post-eligibility transition pattern over time. In the third column, we use only the first three months after losing age eligibility as the post-treatment period. The resulting whole grains coefficient suggests a drop of only 2.6 ounces per month. Effects within the first 6, 9, 12, and 24 months are close to the effect for the entire four years after losing eligibility (ranging from 3.5 to 4.0 ounces less in whole grain products purchased per month). Taken together, the estimates in Table 3.4 suggest that receiving the voucher for whole grain products continues to influence household purchasing patterns even after the vouchers end. The vouchers increase whole grain purchases by about 3.5 ounces per month during eligibility, and three quarters of this effect is reversed within three months after losing eligibility. The rest is undone within six months of losing eligibility.

3.4.3 Effects by Length of Exposure to the New WIC Packages

Finally, we examine whether length of exposure to the new package influences persistence in purchasing patterns after losing eligibility. We use the exact timing of when the new guidelines were implemented in each state to identify the length of time that each household was potentially exposed to these new items before aging out of the WIC program. If strength of habit formation is strictly increasing in length of exposure to incentives, households with longer exposure to the new package would be expected to buy more of the items in the new package for longer after leaving the program.

We restrict attention to households with an age-eligible member within the six months prior to the package change. The identifying variation that we exploit is the timing of when the WIC package was revised relative to the age of the youngest member of the household. For income-eligible households, age of youngest child at the time of the package change determines how long the household was eligible to receive the revised set of vouchers. We compare purchasing patterns after the youngest turns five by length of age eligibility for the new package and income eligibility. Specifically, we estimate the following specification:

$$Y_{ht} = \beta_0 + \beta_1 Exposure_h + \beta_2 IncEligible_h + \beta_3 Exposure_h * IncEligible_h + \gamma X_{ht} + \gamma_t + \epsilon_{ht} \quad (3.3)$$

where $Exposure_h$ measures the number of months that the youngest member of the household was less than 60 months old after the package change, and $IncEligible_h$ indicates whether the household was income eligible within the six months prior to the package change. Y_{ht} is the amount of whole grains or produce, or total food spending, for the household that month. We cannot include household fixed effects, because they would be collinear with the measures of exposure length and income eligibility.

The results are shown in Table 3.5. The sample for the first column includes all households with an age-eligible member within the six months leading up to the package change. Some of the youngest children in these households reached 60 months before the revisions were implemented and so had zero months of exposure to the new vouchers. The point estimates suggest a positive relationship between months of exposure to the new package and whole-wheat purchases and total food spending and a negative relationship with produce purchases. However, none of these coefficients is precisely estimated. In columns two through four, we restrict attention to the first three months, six months, and 12 months after the youngest turned five. The pattern is similar for these subsamples, although the coefficient for whole grains is negative (and smaller) for the six- and 12-month windows. Restricting attention to households with at least one month of age eligibility after the package change produces similar results (see column five). All together, the results do not suggest a strong relationship between length of exposure and persistence in purchasing patterns.

Table 3.5: Impact of Potential Exposure to the Revised WIC Packages on Purchases after Aging Out of Eligibility

	All Periods	Within 3 Months	Within 6 Months	Within 12 Months	All Periods
Whole Grains	0.179	0.054	-0.031	-0.038	0.252
	(0.207)	(0.193)	(0.164)	(0.158)	(0.235)
Produce	-0.027	-0.011	-0.491	-0.405	-0.059
	(0.726)	(0.777)	(0.697)	(0.678)	(0.834)
Food ($)	1.506	0.842	0.616	0.750	1.853
	(0.942)	(1.025)	(0.941)	(0.934)	(1.065)
N	23,233	3,624	6,255	10,844	19,026

Notes: Each row and column displays estimates from a separate regression. These estimates are the coefficients corresponding to the interaction of months of exposure (the amount of time after the WIC package change and before the youngest child reached 60 months) and income eligibility (the household income was below 185% of the FPL in at least one month during the six months prior to the package change). Standard errors are in parentheses and clustered within households. The sample is restricted to households with an age-eligible member within the six months prior to the WIC package change and periods after the youngest child turned five. The second column analyzes the subset of household/month observations that occurred during the three months after the youngest child turned five. The subsamples for the third and fourth columns are analogous. The last column includes only households with at least one month of age eligibility for the new package. Each regression also includes controls for age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school or high school, some college; college graduate omitted), and year, month, and household fixed effects. Whole grains and produce are measured in ounces per month.

3.5 Conclusion

This paper examines whether changing the items included in the WIC bundle impacts what households purchase and whether exposure to these different items results in households continuing to purchase them even after losing WIC eligibility. We exploit two sources of variation to answer these questions. First, we exploit variation in the timing of when the WIC package was implemented in 2009. We use detailed high frequency data on the items that households purchase to compare purchases of specific items before and after the change in the guidelines. Consistent with the goals of the new guidelines we find that households with at least one child eligible for WIC (based on both age and income eligibility) purchased more whole grains after the implementation of the new WIC package. The introduction of the produce voucher did not have a significant effect on produce purchases, consistent with the value of the voucher being inframarginal to household produce budgets.

Second, we exploit variation in the age of the youngest child in the households that are income eligible for WIC at the time of WIC package revision. Some households aged out of WIC just before 2009 and hence had no WIC-induced increase in their exposure to these items. Other households had their youngest child age out in 2010, 2011, etc. thus creating differences across households in the number of years they were exposed to the WIC-induced changes in their purchasing patterns. This creates variation in time exposed to a specific set of items and creates an ideal test of habit formation in purchasing patterns created by a government subsidy. We find limited evidence of a persistent impact on purchases after aging out of eligibility. The persistence in whole-wheat purchases is partial and short-lived, vanishing within six months of losing eligibility. Longer exposure to incentives does not appear to have a substantial effect on strength or length of persistence. Our findings suggest that, for adults, using incentives or vouchers for specific items can be successful in raising consumption of those items while the incentives are in place. However, their efficacy in instilling habits that will outlive program participation is limited.

APPENDIX A
APPENDIX TO CHAPTER 1

Table A.1: Effects on Earnings and Jobs using Alternative Treatment

	Earnings	Jobs
Full sample		
ln(coalres)*(real p coal over 50)	0.001***	0.000**
	(0.000)	(0.000)
first lag	0.001***	0.001***
	(0.000)	(0.000)
second lag	–0.002***	–0.001***
	(0.000)	(0.000)
N	11740	11740
Coal counties		
ln(coalres)*(real p coal over 50)	0.004**	0.002
	(0.002)	(0.001)
first lag	0.007***	0.004***
	(0.002)	(0.001)
second lag	–0.009***	–0.006***
	(0.002)	(0.001)
N	3760	3760

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: Each column displays results from a regression where the outcome variable is an annual log change. Column 1 estimates effects on changes in total county earnings, and column 2 estimates effects for changes in the total number of jobs. The regressor of interest is the natural log of coal reserves times an indicator for peak price years and two lags of this variable. The top panel estimates effects for the full sample of counties in states for which I have coal reserves data. The bottom panel restricts attention to counties with known coal reserves. All specifications include controls for log population in 1970, log change in population from year to year, population per square mile in 1970, unemployment in 1970, share of people below the FPL in 1970, and state by year fixed effects. Standard errors are clustered at the county level.

Table A.2: Effects on Reported Crimes per 1000 People Using Alternative Treatment

	Total	Property	Violent
Full sample			
ln(coalres)*(real p coal over 50)	–0.117**	–0.102**	–0.015
	(0.056)	(0.048)	(0.010)
first lag	0.084	0.065	0.019
	(0.091)	(0.080)	(0.014)
second lag	0.052	0.054	–0.003
	(0.056)	(0.052)	(0.006)
N	8670	8670	8670
Coal counties			
ln(coalres)*(real p coal over 50)	–0.330	–0.263	–0.066
	(0.361)	(0.315)	(0.059)
first lag	0.486	0.355	0.128
	(0.571)	(0.506)	(0.084)
second lag	–0.152	–0.064	–0.085**
	(0.397)	(0.370)	(0.037)
N	2927	2927	2927

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Notes: Each column displays results from a regression where the outcome variable is the annual change in crimes per 1000 people. Column 1 estimate effects on changes in the total number of crimes reported, column 2 estimates effects on changes in the number of property crimes reported, and column 3 estimates effects on changes in the number of violent crimes reported. The regressor of interest is the natural log of coal reserves times an indicator for peak price years and two lags of this variable. The top panel estimates effects for the full sample of counties in states for which I have coal reserves data. The bottom panel restricts attention to counties with known coal reserves. All specifications include controls for log population in 1970, log change in population from year to year, population per square mile in 1970, unemployment in 1970, share of people below the FPL in 1970, and state by year fixed effects. Standard errors are clustered at the county level.

APPENDIX B
APPENDIX TO CHAPTER 2

Table B.1: Five Strongest Selected Instruments for IVNM

			Weight
Panel A: Felonies			
Crime Type			0.111
Crime Type	×	Number of Counts	0.076
Crime Type	×	Prior Felony Convictions	0.062
Crime Type	×	First Time Offender	0.052
Crime Type	×	Prior Felony Arrests	0.045
Panel A: Misdemeanors			
Crime Type			0.067
Crime Type	×	Number of Counts	0.052
Crime Type	×	Prior Misdemeanor Convictions	0.042
Crime Type	×	First Time Offender	0.036
Crime Type	×	Prior Felony Convictions	0.035

Notes: This tables reports the five strongest predictors of pretrial detention selected by Lasso from the set of all candidate interactions. Due to the differing degree of variance for each interaction, the instruments are each normalized to mean zero and standard deviation one (Mueller-Smith, 2014). Crime type interacted with a defendant's criminal history are the characteristics over which judges tend to exhibit the largest differences.

Table B.2: Test for Judge Random Assignment by Covariate

	Felonies	Misdemeanors
Outcomes		
Detained Pretrial	10.36	19.29
Released on Own Recognizance	19.80	18.35
Set Bail	19.18	30.90
Remanded	3.22	6.63
Demographics		
Age	1.47	1.30
Female	1.18	2.73
White	1.19	1.20
Black	1.45	1.58
Hispanic	1.56	1.39
Criminal History		
Sex Offender	0.96	1.07
First Time Offender	1.45	1.19
Prior Felony Arrests	1.38	1.18
Prior Misdemeanor Arrests	1.63	1.24
Prior Felony Convictions	1.30	1.05
Prior Misdemeanor Convictions	1.58	1.17
Number of Counts	1.62	0.82

Notes: Each cell in the table reports the F statistic when testing for the joint significance of the judge binary variables for the indicated outcome variable conditional on courthouse by time of year by arraignment shift fixed effects. The first column is for felonies and the second column is for misdemeanors.

Table B.3: Test for Judge Random Assignment using Judge Leave-Out Mean

	Felonies		Misdemeanors	
	Predicted	Actual	Predicted	Actual
Detained	0.0062	0.0601	0.0023	0.0248
	(0.0006)	(0.0013)	(0.0002)	(0.0005)
Released on Own Recognizance	-0.0064	-0.0592	-0.0008	-0.0168
	(0.0007)	(0.0014)	(0.0006)	(0.0009)
Set Bail	0.0051	0.0542	0.0028	0.0320
	(0.0007)	(0.0014)	(0.0002)	(0.0005)
Remanded	0.0014	0.0064	0.0000	0.0003
	(0.0003)	(0.0005)	(0.0000)	(0.0001)

Notes: Each cell in the table reports the coefficient on the leave-out mean for how much a judge deviates from the full-sample average detention rate for each Uniform Crime Reporting crime type code (z_{ijc}) when normalized, along with its standard error. The coefficient on the normalized z_{ijc} comes from a regression of the outcome variable shown in each row (either predicted or actual) on the normalized leave-out mean conditional on courthouse by time of year by arraignment shift fixed effects. For the first and third columns, the outcome variables are the predicted values for the indicated variables in the variable column that are obtained from a regression of the indicated variable on demographic, criminal history, and courthouse by time of year by arraignment shift fixed effects controls. The second and fourth columns use the actual (not predicted) indicated variables as the outcome variables. Standard errors clustered at the shift level are reported.

Table B.4: Effect of Detention on Felony Case Outcomes Using Judge FE IV

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conviction						
Detained	0.098***	0.100***	0.101***	0.116***	0.148***	0.143***
	(0.024)	(0.024)	(0.023)	(0.014)	(0.011)	(0.011)
Panel B: Plead Guilty						
Detained	0.111***	0.112***	0.118***	0.122***	0.179***	0.164***
	(0.025)	(0.025)	(0.024)	(0.015)	(0.012)	(0.011)
Panel C: Minimum Sentence Length (Days)						
Detained	28	19	-3	110***	109***	114***
	(31)	(31)	(27)	(20)	(15)	(14)
Time and Courtroom FE	X	X	X	X	X	X
Demographics		X	X	X	X	X
Criminal History			X	X	X	X
Judge by Criminal History				X	X	X
Judge by Top 5 Crimes					X	X
Judge by Race and Gender						X
F Statistic	9.96	11.27	11.28	5.98	4.49	3.54
Number of Instruments	212	212	212	843	1,891	2,728
Observations	244,221	244,221	244,221	244,221	244,221	244,221

Notes: Table B.4 reports estimates of the effect of pretrial detention on case outcomes for felonies using instrumental variables. The first three columns use the judge binary variables as instruments for being detained with varying controls. The instruments used for column 4 are judge binary variables interacted with the number of prior felony convictions, prior misdemeanor convictions, and a binary variable for being a first time offender. Column 5 adds the interactions of the judge binary variables and the binary variables for the top five crime types as instruments. Column 6 adds the interactions of the judge binary variables and the binary variables for black, white, Hispanic, and female as instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.5: Effect of Detention on Misdemeanor Case Outcomes Using Judge FE IV

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conviction						
Detained	0.174***	0.165***	0.152***	0.091***	0.086***	0.080***
	(0.028)	(0.027)	(0.025)	(0.010)	(0.009)	(0.009)
Panel B: Plead Guilty						
Detained	0.203***	0.195***	0.184***	0.099***	0.088***	0.080***
	(0.028)	(0.028)	(0.025)	(0.010)	(0.009)	(0.009)
Panel C: Minimum Sentence Length (Days)						
Detained	-6	-7	-8*	21***	23***	22***
	(5)	(5)	(5)	(2)	(2)	(2)
Time and Courtroom FE	X	X	X	X	X	X
Demographics		X	X	X	X	X
Criminal History			X	X	X	X
Judge by Criminal History				X	X	X
Judge by Top 5 Crimes					X	X
Judge by Race and Gender						X
F Statistic	14.20	14.35	15.34	26.98	14.01	10.23
Number of Instruments	212	212	212	845	1,900	2,744
Observations	726,785	726,780	726,780	726,780	726,780	726,780

Notes: Table B.5 reports estimates of the effect of pretrial detention on case outcomes for misdemeanors using instrumental variables. The first three columns use the judge binary variables as instruments for being detained with varying controls. The instruments used for column 4 are judge binary variables interacted with the number of prior felony convictions, prior misdemeanor convictions, and a binary variable for being a first-time offender. Column 5 adds the interactions of the judge binary variables and the binary variables for the top five crime types as instruments. Column 6 adds the interactions of the judge binary variables and the binary variables for black, white, Hispanic, and female as instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.6: Effect of Detention for Different Misdemeanor Samples

	All Misdemeanors			Exclude Dismissed			Exclude Plead Guilty			Exclude Both		
	OLS	IV	IVNM	OLS	IV	IVNM	OLS	IV	IVNM	OLS	IV	IVNM
Panel A: Conviction												
Detained	0.062***	0.149***	0.071***	0.052***	0.000	0.043***	0.124***	0.226***	0.098***	0.100***	0.077***	0.074***
	(0.002)	(0.022)	(0.006)	(0.002)	(0.020)	(0.005)	(0.002)	(0.022)	(0.006)	(0.002)	(0.020)	(0.005)
Panel B: Plead Guilty												
Detained	0.051***	0.110***	0.056***	0.040***	-0.040**	0.026***	0.120***	0.208***	0.098***	0.096***	0.058***	0.071***
	(0.002)	(0.022)	(0.006)	(0.002)	(0.020)	(0.005)	(0.002)	(0.022)	(0.006)	(0.002)	(0.020)	(0.005)
Panel C: Minimum Sentence Length (Days)												
Detained	17***	56***	30***	17***	54***	29***	12***	36***	19***	11***	35***	17***
	(0)	(4)	(1)	(0)	(4)	(1)	(0)	(4)	(1)	(0)	(4)	(1)
F Statistic	-	3,113	1,108	-	2,982	1,113	-	3,304	1,288	-	3,215	1,388
Number of Instruments	-	1	26	-	1	26	-	1	25	-	1	23
Observations	1,106,373	1,103,832	1,103,148	873,539	870,975	870,291	961,587	959,093	958,407	728,750	726,167	725,480

Notes: Table B.6 reports estimates of the effect of pretrial detention on case outcomes for misdemeanors from different misdemeanor samples. The first three columns include all misdemeanor arraignments. The second three columns exclude arraignments for which the case was dismissed at the arraignment. The third three columns exclude arraignments for which the defendant plead guilty at the arraignment. The fourth three columns exclude arraignments for which the case was either dismissed at the arraignment or the defendant plead guilty at the arraignment. The OLS columns use the same specification as column 8 of Table 2.5 and report estimates from an OLS specification with all the baseline controls included. The IV columns use the same specification as column 3 of Table 2.7 and report estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. The IVNM columns use the same specification as column 6 of Table 2.7 and report estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.7: Misdemeanor IV Estimates for Cases Types with Low Levels of Adjudication

	Misdemeanors					
	IV			IVNM		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conviction						
Detained	0.227***	0.306***	0.259***	0.214***	0.204***	0.142***
	(0.079)	(0.080)	(0.080)	(0.012)	(0.012)	(0.014)
Panel B: Plead Guilty						
Detained	0.238***	0.315***	0.272***	0.217***	0.206***	0.142***
	(0.079)	(0.079)	(0.079)	(0.012)	(0.012)	(0.014)
Panel C: Minimum Sentence Length (Days)						
Detained	6	4	0	25***	24***	15***
	(15)	(15)	(16)	(2)	(2)	(3)
Time and Courtroom FE	X	X	X	X	X	X
Demographics		X	X		X	X
Criminal History			X			X
F Statistic	380	360	379	201	213	237
Number of Instruments	1	1	1	23	23	23
Observations	215,755	215,751	215,751	215,727	215,727	215,727

Notes: Table B.7 reports estimates of the effect of pretrial detention on case outcomes for DUI, simple assault, and aggravated assault misdemeanors using instrumental variables. The first three columns use the judge-level variation in the propensity to detain pretrial for a given crime type as an instrument for being detained with varying controls. The last three columns use the set of Lasso selected instruments from the full set of constructed candidate instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.8: Effects on Conviction by Predicted Incarceration Length for Felonies

	Felonies		
Predicted Incarceration Length	OLS	IV	IVNM
0 Days	0.108***	0.125**	0.102***
	(0.006)	(0.060)	(0.014)
1-290 Days	0.142***	0.119***	0.114***
	(0.004)	(0.030)	(0.008)
290+ Days	0.168***	0.183***	0.155***
	(0.004)	(0.028)	(0.008)

Notes: Table B.8 reports estimates of the effect of pretrial detention on conviction for felonies by the predicted length of incarceration if convicted. Column 1 uses the same specification as column 4 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Column 2 uses the same specification as column 3 of Table 2.6 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Column 3 uses the same specification as column 6 of Table 2.6 and reports estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.9: Effects on Conviction by Predicted Incarceration Length for Misdemeanors

	Misdemeanors		
Predicted Incarceration Length	OLS	IV	IVNM
0 Days	0.045***	-0.496**	0.052*
	(0.007)	(0.204)	(0.029)
1-11 Days	0.110***	0.188***	0.052***
	(0.004)	(0.067)	(0.015)
11+ Days	0.103***	0.075***	0.080***
	(0.002)	(0.017)	(0.006)

Notes: Table B.9 reports estimates of the effect of pretrial detention on conviction for misdemeanors by the predicted length of incarceration if convicted. Column 1 uses the same specification as column 8 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Column 2 uses the same specification as column 3 of Table 2.7 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Column 3 uses the same specification as column 6 of Table 2.7 and reports estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.10: Effects on Conviction by Top 10 Crime Type

	Felonies		
Crime Type	OLS	IV	IVNM
Robbery	0.134***	0.155***	0.126***
	(0.005)	(0.040)	(0.014)
Drug Sales	0.206***	0.141*	0.232***
	(0.007)	(0.073)	(0.021)
Aggravated Assault	0.121***	0.066	0.134***
	(0.007)	(0.098)	(0.017)
Drug Possession	0.184***	0.356***	0.200***
	(0.008)	(0.121)	(0.024)
Larceny	0.146***	0.104	0.146***
	(0.009)	(0.286)	(0.024)
Burglary	0.174***	0.166	0.169***
	(0.010)	(0.112)	(0.023)
Simple Assault	0.159***	0.400	0.155***
	(0.012)	(0.392)	(0.025)
Weapon	0.121***	0.146	0.126***
	(0.012)	(0.275)	(0.026)
Forgery	0.120***	0.308	0.114***
	(0.013)	(0.564)	(0.025)
Other	0.091***	-0.008	0.096***
	(0.007)	(0.077)	(0.010)

Notes: Table B.10 reports estimates of the effect of pretrial detention on conviction for felonies by the top 10 felony crime types. Each row indicates the type of crime the sample is restricted to. Column 1 uses the same specification as column 4 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Column 2 uses the same specification as column 3 of Table 2.6 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Column 3 uses the same specification as column 6 of Table 2.6 and reports estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.11: Effects on Conviction by Top 10 Crime Type

	Misdemeanor		
Crime Type	OLS	IV	IVNM
Drug Possession	0.019***	-0.032	0.013
	(0.003)	(0.029)	(0.010)
Simple Assault	0.202***	0.163*	0.151***
	(0.004)	(0.086)	(0.019)
Fraud	-0.027***	0.191*	-0.023
	(0.006)	(0.098)	(0.017)
Larceny	0.060***	0.066	0.019
	(0.005)	(0.057)	(0.014)
Weapon	-0.001	0.836	0.018
	(0.011)	(0.723)	(0.028)
Criminal Mischief	0.022	0.597	0.012
	(0.025)	(1.067)	(0.038)
DUI	0.188***	0.121	0.179***
	(0.011)	(0.305)	(0.026)
Aggravated Assault	0.236***	-0.097	0.160***
	(0.011)	(0.292)	(0.027)
Drug Sale	0.064***	0.568**	0.073***
	(0.014)	(0.255)	(0.026)
Other	0.102***	0.105	0.093***
	(0.004)	(0.074)	(0.010)

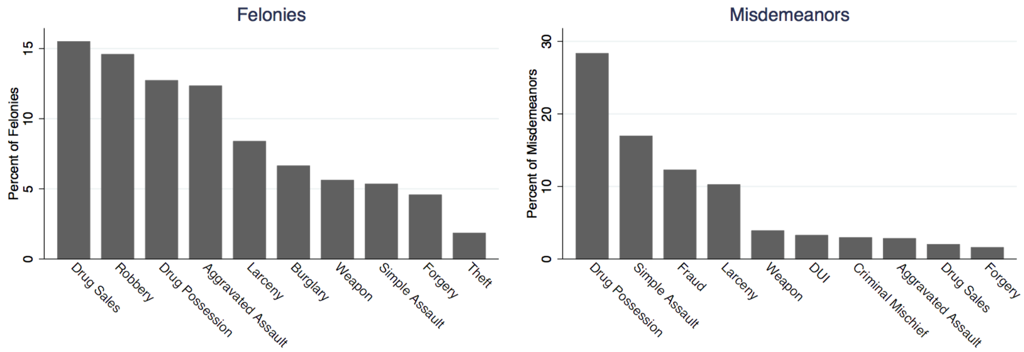
Notes: Table B.11 reports estimates of the effect of pretrial detention on conviction for misdemeanors by the top 10 misdemeanor crime types. Each row indicates the type of crime the sample is restricted to. Column 1 uses the same specification as column 8 of Table 2.5 and reports estimates from an OLS specification with all the baseline controls included. Column 2 uses the same specification as column 3 of Table 2.7 and reports estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Column 3 uses the same specification as column 6 of Table 2.7 and reports estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B.12: Effects by Predicted Incarceration Length and Gender for Felonies

	Men			Women		
Predicted Incarceration Length	OLS	IV	IVNM	OLS	IV	IVNM
0 Days	0.100***	0.133*	0.107***	0.122***	0.149	0.067**
	(0.007)	(0.068)	(0.016)	(0.014)	(0.153)	(0.032)
1-290 Days	0.139***	0.113***	0.113***	0.151***	0.142**	0.130***
	(0.004)	(0.034)	(0.009)	(0.012)	(0.071)	(0.026)
290+ Days	0.163***	0.182***	0.153***	0.194***	0.123	0.133***
	(0.004)	(0.030)	(0.009)	(0.022)	(0.127)	(0.039)

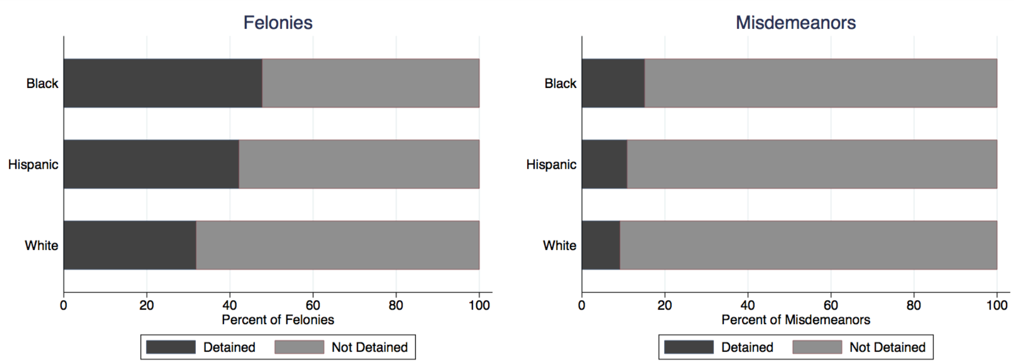
Notes: Table B.12 reports estimates of the effect of pretrial detention on conviction for felonies by the predicted length of incarceration if convicted and by gender. Columns 1 and 4 use the same specification as column 4 of Table 2.5 and report estimates from an OLS specification with all the baseline controls included. Columns 2 and 5 use the same specification as column 3 of Table 2.6 and report estimates from an IV estimation that uses judge-level variation in the propensity to detain pretrial for a given crime type as an instrument. Columns 3 and 6 use the same specification as column 6 of Table 2.6 and report estimates from an IV estimation that uses the set of Lasso selected instruments. Standard errors clustered at the shift level are reported. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Figure B.1: Most Common Types of Offenses



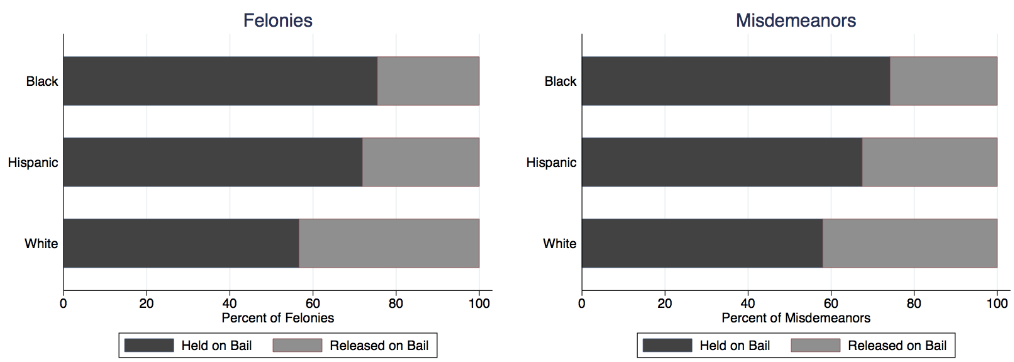
Notes: Each graph shows the most common types of offenses. For each bar, the y-axis shows what percent of offenses were of the indicated type. The first graph is for felonies and the second is for misdemeanors.

Figure B.2: Fraction Detained by Race and Ethnicity



Notes: Each graph shows the percent of defendants who were detained for black, Hispanic, and white defendants. The first part of each bar shows the percent of defendants detained. The second part of each bar shows the percent of defendants not detained.

Figure B.3: Fraction Held on Bail by Race and Ethnicity



Notes: Each graph shows the percent of black, Hispanic, and white defendants that were held on bail when bail was set. The first part of each bar shows the percent of defendants held on bail. The second part of each bar shows the percent of defendants released on bail.

APPENDIX C
APPENDIX TO CHAPTER 3

Table C.1: Difference-in-Differences Estimates of the Impact of the WIC Package Change for Additional Food Purchases

	All Milk	Low-fat Milk	Whole Milk	Pct. Low-Fat Milk	All Cereal	Whole Grain Cereals	Sugar Cereals	Cheese	Eggs	Juice	Fish	Produce ($)
Panel A: Sample is Households With At Least One Age-Eligible Child												
Income-Eligible X After	16.700	30.890	-14.190	4.636	0.744	1.022	-0.400	1.660	0.433	5.531	-0.202	0.424
	(17.090)	(16.900)	(9.874)	(1.305)	(1.637)	(0.779)	(1.079)	(1.234)	(1.537)	(9.823)	(0.386)	(0.555)
Income-Eligible	-32.500	-45.230	12.730	-3.295	-3.994	-1.671	-1.791	-3.620	-1.668	-16.830	0.182	-0.485
	(16.490)	(17.240)	(12.430)	(1.311)	(1.413)	(0.632)	(1.020)	(1.094)	(1.385)	(10.360)	(0.341)	(0.456)
After the Package Change	-19.300	-20.850	1.549	-1.404	-1.265	-1.835	0.511	-1.927	-0.741	1.217	0.147	-1.189
	(9.416)	(9.363)	(6.370)	(0.802)	(1.176)	(0.559)	(0.761)	(0.848)	(1.013)	(6.298)	(0.274)	(0.348)
Pre-Revision Mean	573.34	408.95	164.39	58.58	64.09	16.49	33.31	49.33	51.79	323.13	8.10	16.86
Observations	134,624	134,624	134,624	134,624	180,796	180,796	180,796	180,796	180,796	180,796	180,796	180,796
Panel B: Sample is Households Without Any Age-Eligible Children												
Income-Eligible X After	15.370	9.878	5.496	3.297	0.564	0.008	-0.450	2.006	0.966	-1.909	-0.722	1.048
	(16.310)	(15.110)	(10.140)	(1.376)	(1.998)	(0.872)	(1.508)	(1.495)	(2.002)	(12.210)	(0.533)	(0.592)
Income-Eligible	-27.840	-16.060	-11.780	-2.974	-0.519	-0.383	1.029	-3.031	1.435	7.396	-0.172	-0.051
	(15.570)	(15.200)	(10.070)	(1.571)	(1.889)	(0.884)	(1.481)	(1.539)	(1.962)	(12.110)	(0.525)	(0.546)
After the Package Change	-10.290	-8.888	-1.404	-1.372	-0.002	-0.929	0.923	-1.110	-1.675	9.148	0.810	-0.246
	(9.582)	(8.940)	(5.058)	(0.821)	(1.473)	(0.776)	(0.978)	(1.126)	(1.468)	(7.784)	(0.390)	(0.400)
Pre-Revision Mean	533.03	399.02	134.02	61.58	66.91	15.39	38.51	44.98	47.54	304.73	8.14	16.54
Observations	79,444	79,444	79,444	79,444	93,035	93,035	93,035	93,035	93,035	93,035	93,035	93,035

Notes: Each column displays estimates from separate regressions. In both panels, the sample is restricted to households in which the youngest child is at least 12 months of age. For milk, the sample is restricted to households in which the youngest child is at least 24 months of age, since the milk items in the WIC package differ for children below 24 months from the package for children at least 24 months of age. Additional variables included, but not shown, are age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school, high school, some college; college graduate omitted), and year, month, and household fixed effects. All units are in ounces except as specified. The pre-revision means are the means prior to the package change for income-eligible households for each sample. Fish is not included as part of the WIC package for children, either before or after the package change, but is included as part of the package for fully breastfeeding mothers. Cereal includes all hot and cold cereals, not just whole grain and sugar cereals. Standard errors are in parentheses and clustered within households.

Table C.2: The Relationship between the WIC Package Change and Household Characteristics

	Household Income	Household Size	Black	Hispanic	Other Race/ Ethnicity	Married	High School or Less	Some College	College Graduate	Age of the Youngest Child
Income-Eligible X After	-0.738	-0.059	0.000	0.004	-0.003	-0.023	0.004	-0.003	-0.001	-0.233
	(0.905)	(0.039)	(0.004)	(0.004)	(0.006)	(0.009)	(0.008)	(0.012)	(0.010)	(0.837)
Income-Eligible	-21.020	0.416	0.003	0.001	-0.005	0.019	0.002	0.001	-0.003	-1.481
	(0.988)	(0.038)	(0.003)	(0.004)	(0.005)	(0.007)	(0.009)	(0.012)	(0.009)	(0.669)
After the Package Change	0.296	0.020	0.001	0.000	0.001	0.007	0.000	-0.005	0.004	-0.074
	(0.445)	(0.016)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.325)
Pre-Revision Mean	28.97	4.89	0.12	0.11	0.05	0.75	0.27	0.35	0.37	34.91
Observations	180,796	180,796	180,796	180,796	180,796	180,796	180,796	180,796	180,796	180,796

Notes: Each column displays estimates from separate regressions, where the outcome variables are denoted by the column heading. The sample is restricted to households were the youngest child is at least 12 months of age and there is at least one age-eligible child in the household. Standard errors are in parentheses and clustered within households. Additional variables included, but not shown, are age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school or high school, some college; college graduate omitted), and year, month, and household fixed effects. The pre-revision means are the means prior to the package change for income-eligible households.

Table C.3: Difference-in-Differences Estimates of the Impact of Aging out of WIC Eligibility for Additional Food Purchases

		After the Package Change			
	Before the Package Change	All Periods	First 3 Months	First 6 Months	First 12 Months
All Milk	10.790	5.990	15.400	10.230	10.760
	(14.990)	(12.810)	(13.390)	(12.820)	(13.560)
Low-Fat Milk	29.600	-10.750	7.103	-2.773	-6.914
	(14.070)	(13.340)	(12.680)	(12.470)	(14.110)
Whole Milk	-18.810	16.740	8.293	13.000	17.670
	(9.604)	(10.560)	(7.572)	(8.551)	(11.530)
Pct. Low-Fat Milk	1.998	-1.554	0.185	-0.679	-1.575
	(1.199)	(1.058)	(1.223)	(1.162)	(1.147)
All Cereals	0.676	0.169	0.088	-1.327	-0.963
	(1.730)	(1.607)	(1.943)	(1.824)	(1.779)
Whole Grain Cereal	0.082	-0.969	-1.628	-1.870	-1.830
	(0.763)	(0.736)	(0.950)	(0.858)	(0.797)
Sugar Cereals	2.239	1.841	1.962	1.250	1.630
	(1.370)	(1.091)	(1.370)	(1.266)	(1.261)
Cheese	-2.317	-0.805	-1.503	-2.355	-1.042
	(1.108)	(1.170)	(1.420)	(1.222)	(1.248)
Eggs	-0.658	-1.010	-1.714	-2.585	-1.990
	(1.817)	(1.513)	(1.848)	(1.613)	(1.573)
Juice	-4.249	-1.333	-8.749	-7.532	-1.954
	(11.250)	(9.310)	(10.860)	(10.250)	(10.070)
Fish	-0.076	-0.679	-0.206	-0.422	-0.338
	(0.500)	(0.341)	(0.472)	(0.424)	(0.407)
Produce ($)	-0.722	-2.099	-1.538	-1.724	-1.485
	(0.373)	(0.672)	(0.730)	(0.693)	(0.708)
N	133,882	139,949	91,243	95,972	104,404

Notes: Each row and column displays estimates from separate regressions. These estimates are the coefficients corresponding to the interaction of income-eligibility (under 185% FPL) and having the youngest member of the household at least 60 months old (no longer age-eligible). Standard errors are in parentheses and clustered within households. The sample is restricted to households where the youngest child is within 48 months of turning 60 months of age. Additional variables included, but not shown, are age of the youngest in the household, household income, household size, race/ethnicity (black, Hispanic, and other race/ethnicity; white omitted), married, educational attainment (less than high school or high school, some college; college graduate omitted), and year, month, and household fixed effects. All units are in ounces except as specified.

REFERENCES

Aizer, A. and Doyle, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, 130(2):759–803.

Anderson, D. A. (1999). The aggregate burden of crime. *The Journal of Law and Economics*, 42(2):611–642.

Andreyeva, T. and Luedicke, J. (2013). Federal food package revisions: effects on purchases of whole-grain products. *American journal of preventive medicine*, 45(4):422–429.

Angrist, J. D. and Kugler, A. D. (2008). Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia. *The Review of Economics and Statistics*, 90(2):191–215.

Ares, C. E., Rankin, A., and Sturz, H. (1963). The manhattan bail project: An interim report on the use of pre-trial parole. *NYU L aw Review*, 38:67.

Arezki, R. and Van der Ploeg, F. (2011). Do natural resources depress income per capita? *Review of Development Economics*, 15(3):504–521.

Auty, R. M. (2001). The political economy of resource-driven growth. *European Economic Review*, 45(4):839–846.

Barry, J. (2014). *Criminal Court of the City of New York: 2014 Annual Report*. Office of the Chief Clerk of New York City Criminal Court.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

Bell, B., Fasani, F., and Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, 21(3):1278–1290.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429.

Belot, M., James, J., and Nolen, P. (2013). Changing eating habits-a field experiment in primary schools. Working Paper, University of Edinburgh.

Birch, L. L. and Marlin, D. W. (1982). I don't like it; I never tried it: Effects of exposure on two-year-old children's food preferences. *Appetite*, 3(4):353–360.

Bitler, M. P., Currie, J., and Scholz, J. K. (2003). WIC eligibility and participation. *Journal of Human Resources*, pages 1139–1179.

Black, D., Daniel, K., and Sanders, S. (2002). The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust. *The American Economic Review*, 92(1):27.

Black, D., McKinnish, T., and Sanders, S. (2005). The economic impact of the coal boom and bust. *The Economic Journal*, 115(503):449–476.

Black, D. A., McKinnish, T. G., and Sanders, S. G. (2003). Does the availability of high-wage jobs for low-skilled men affect welfare expenditures? Evidence from shocks to the steel and coal industries. *Journal of Public Economics*, 87(9):1921–1942.

Brown, A., Berryhill Jr, H. L., Taylor, D. A., and Trumbull, J. V. (1952). *Coal resources of Virginia*. United States Geological Survey.

Cascio, E. U. and Narayan, A. (2015). Who needs a fracking education? the educational response to low-skill biased technological change. NBER Working Paper No. 21359.

Chalfin, A. (2014). What is the contribution of mexican immigration to us crime rates? Evidence from rainfall shocks in mexico. *American Law and Economics Review*, 16(1):220–268.

Charness, G. and Gneezy, U. (2009). Incentives to exercise. *Econometrica*, 77(3):909–931.

Cohee, G. V., Burns, R., Brown, A., Brant, R., and Wright, D. (1950). *Coal resources of Michigan*. United States Geological Survey.

Cosgrove, B. M., LaFave, D. R., Dissanayake, S. T., and Donihue, M. R. (2015). The economic impact of shale gas development: A natural experiment along the New York/Pennsylvania border. *Agricultural and Resource Economics Review*, 44:20–39.

Dobbie, W., Goldin, J., and Yang, C. (2016). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. NBER Working Paper No. 22511.

Feyrer, J., Mansur, E. T., and Sacerdote, B. (2015). Geographic dispersion of economic shocks: Evidence from the fracking revolution. NBER Working Paper No. 21624.

Freeman, R. B. (2000). Disadvantaged young men and crime. In *Youth employment and joblessness in advanced countries*, pages 215–246. University of Chicago Press.

GAO (2013). WIC program: Improved oversight of income eligibility determination needed. Technical report, GAO-13-290.

Gould, E. D., Weinberg, B. A., and Mustard, D. B. (2002). Crime rates and local labor market opportunities in the United States: 1979–1997. *Review of Economics and Statistics*, 84(1):45–61.

Gupta, A., Hansman, C., and Frenchman, E. (2016). The heavy costs of high bail: Evidence from judge randomization. *Columbia Law and Economics Working Paper*, (531).

Harding, M., Leibtag, E., and Lovenheim, M. F. (2012). The heterogeneous geographic and socioeconomic incidence of cigarette taxes: Evidence from Nielsen Homescan Data. *American Economic Journal: Economic Policy*, 4(4):169–198.

Holder, E. (2011). Attorney general eric holder speaks at the national symposium on pretrial justice. https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-symposium-pretrial-justice.

Hoynes, H. W. and Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–139.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

James, P. (1984). The future of coal. In *The Future of Coal*, pages 228–241. Springer.

Keenan, C. T. (2013). *We need more bail reform*. Pretrial Services Agency for the District of Columbia.

Kellough, G. and Wortley, S. (2002). Remand for plea. bail decisions and plea bargaining as commensurate decisions. *British Journal of Criminology*, 42(1):186–210.

Kelly, M. (2000). Inequality and crime. *Review of Economics and Statistics*, 82(4):530–539.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *The American Economic Review*, 96(3):863.

Komarek, T. (2014). Crime and natural resource booms: Evidence from unconventional natural gas production. Working Paper.

Kreider, B., Pepper, J. V., and Roy, M. (2016). Identifying the effects of WIC on food insecurity among infants and children. *Southern Economic Journal*.

Leiber, M. J. and Fox, K. C. (2005). Race and the impact of detention on juvenile justice decision making. *Crime & Delinquency*, 51(4):470–497.

Lin, M.-J. (2008). Does unemployment increase crime? Evidence from us data 1974–2000. *Journal of Human Resources*, 43(2):413–436.

List, J. A. and Samek, A. S. (2015). The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption. *Journal of Health Economics*, 39:135–146.

Livingston, G. (2013). *The rise of single fathers: A ninefold increase since 1960.* Pew Research Center.

Loewenstein, G., Price, J., and Volpp, K. (2016). Habit formation in children: Evidence from incentives for healthy eating. *Journal of Health Economics*, 45:47–54.

Machin, S. and Meghir, C. (2004). Crime and economic incentives. *Journal of Human Resources*, 39(4):958–979.

Maltz, M. D. and Targonski, J. (2002). A note on the use of county-level UCR data. *Journal of Quantitative Criminology*, 18(3):297–318.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. Working Paper.

Nunziata, L. (2015). Immigration and crime: Evidence from victimization data. *Journal of Population Economics*, 28(3):697–736.

Phillips, M. T. (2008). *Pretrial detention and case outcomes, part 2: Felony cases.* CJA, New York City Criminal Justice Agency.

Phillips, M. T. (2012). Decade of bail research in new york city. *New York City Criminal Justice Agency.*

Rabin, M. (2013). Healthy habits: Some thoughts on the role of public policy in healthful eating and exercise under limited rationality. In Oliver, A., editor, *Behavioral Public Policy*. Cambridge University Press.

Rankin, A. (1964). The effect of pretrial detention. *NYU Law Review*, 39:641.

Raphael, S. and Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *Journal of Law and Economics*, 44(1):259–283.

Reaves, B. A. (2013). Felony defendants in large urban counties, 2009 - statistical tables. Technical Report NCJ 243777, Bureau of Justice Statistics.

Royer, H., Stehr, M., and Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, 7(3):51–84.

Spencer, F. D. (1953). *Coal resources of Indiana*. United States Geological Survey.

Stevenson, M. (2016). Distortion of justice. *Working Paper.*

USDA (2015a). Frequently asked questions about wic. www.fns.usda.gov/wic/frequently-asked-questions-about-wic. Retrieved November 5, 2015.

USDA (2015b). Wic program participation and costs. http://www.fns.usda.gov/sites/default/files /pd/wisummary.pdf. Retrieved November 9, 2015.

Van der Ploeg, F. (2011). Natural resources: curse or blessing? *Journal of Economic Literature*, 49(2):366–420.

Wagner, P. (2015). Jails matter. But who is listening? Prison Policy Initiative.

Wagner, P. and Sakala, L. (2014). Mass incarceration: The whole pie. *Northampton, MA: Prison Policy Initiative*.

Whaley, S. E., Ritchie, L. D., Spector, P., and Gomez, J. (2012). Revised wic food package improves diets of wic families. *Journal of Nutrition Education and Behavior*, 44(3):204–209.

Williams, M. R. (2003). The effect of pretrial detention on imprisonment decisions. *Criminal Justice Review*, 28(2):299–316.

Zenk, S. N., Powell, L. M., Odoms-Young, A. M., Krauss, R., Fitzgibbon, M. L., Block, D., and Campbell, R. T. (2014). Impact of the revised special supplemental nutrition program for women, infants, and children (wic) food package policy on fruit and vegetable prices. *Journal of the Academy of Nutrition and Dietetics*, 114(2):288–296.