Deviance and Government Capacity

Jim Bisbee

New York University

jhb362@nyu.edu

Abstract

The decline in crime rates over the last three decades has enlivened the debate over the preferred methods of combating deviant behavior. Simple regressions of crime on income have documented a strong negative relationship, suggesting that crime is a problem of the poor. However, this research has been unable to adjudicate between the competing theories for this negative relationship. One of the most entrenched debates is over the mediating effects of the state. Do changes in income cause declines in deviant behavior because the government is better able to deter crime with expanded budgets? Or is the causal effect primarily the result of increased opportunity costs? I test these competing theories on county-level crime data from the United States between 1982 and 2007, using trade shocks as a source of exogenous variation in income. I conclude that the majority of the income-crime relationship is explained by the direct effects of opportunity costs. However, there is evidence of a small but significant mediation effect in the form of government expenditures on police.

I. Introduction

Is the negative relationship between wealth and crime primarily the product of an opportunity cost channel or a state capacity channel? This question captures the current status of a growing body of literature attempting to understand how and why the rule of law breaks down. Criminal or "deviant" behavior is a problem faced by all societies and one for whom poorer groups are systematically more affected. This pattern persists both intra- and inter-nationally.

However, there remains an unresolved theoretical debate over whether the income-crime relationship is fundamentally one of opportunity costs or one of government capacity. Put more starkly, do individuals turn to crime when alternative means of obtaining resources are lost to them, or will individuals always predate on each other without formal structures designed to restrict this behavior? These competing theories have been popularized most recently in the literature on conflict which has been unable to disentangle their effects.

This paper tests the relative strength of the competing theories on panel data of US crime disaggregated to the county level from 1982 to 2007. I find that a 1% increase in net residential income causes a decline in property crimes by 6.4 offenses per 10,000 individuals, of which only 0.3 offenses are attributable to the government capacity channel. This result supports the opportunity cost theories in the literature and has implications for policy solutions targeting crime.

To address causality concerns, I use an instrumental variable (IV) analysis borrowed from the labor economics literature that exploits exogenous variation in income generated by trade shocks (David et al. [25], Feler and Senses [30]). The instrument calculates import penetration weights by industry and uses changes in these measures to predict changes in income. Under the assumption that trade shocks are orthogonal to criminal behavior except through their effects on income, this instrument ensures that estimates of the income-crime relationship are causally identified.

This paper proceeds as follows. Section II motivates the research by placing it in the context of the existing literature. Section III sketches a simple diagram of the empirical predictions I will test. Section IV introduces the empirical methods and summarizes the data. Section V presents the main empirical results. Section VI examines the robustness of the results. Section VII discusses the implications for the existing literature as well as policy. Section VIII concludes.

II. Literature Review

Criminology theories frame the decision to commit a crime as the product of a cost-benefit analysis in which a citizen chooses strategies based on legal income and the returns to criminal activity.[1] Two competing predictions fall out of these models, depending on whether the citizen experiences a decrease in relative or absolute income. A decline in relative income (growing inequality which can occur despite all citizens growing wealthier) prompts criminal behavior by expanding the range of attractive targets for predation (Cohen and Felson [21], Land [42], Brush [17]). Competing theories predict that an absolute increase income reduces criminal behavior via increased opportunity costs (Trumbull [61], Doyle et al. [27], McCall et al. [44]). The criminology literature has resolved the empirical debate in favor of a negative causal relationship (Bignon et al. [15]) when incomes rise, even if the rise is accompanied by growing inequality.

A different set criminological theories argue that the citizen chooses strategies based on a number of factors related to government capacity, including deterrence, socialization, and *anomie*.[2] The deterrence literature posits that a greater likelihood of conviction and harsher punishments combine to deter deviant behavior (Becker [10] Ehrlich [28]). Meanwhile social disorganization theories argue that public spending on education reduces crime by instilling pro-social values in youths who would otherwise succumb to subcultures of violence.[3] Finally, strain theory contends that publicly funded welfare programs are vital for combating the crime-inducing effects of *anomie* (Messner [47, 48, 49], Chamlin and Cochran [20], Colvin et al. [22]).

The criminology literature is faced with a theoretical debate over the channels by which income negatively affects deviant behavior. This debate hinges on whether the effect is transmitted primarily via opportunity costs or via government capacity. Within the government capacity channel, there remain competing views on the relative effects of hard versus soft deterrence. Criminology's discussion of government spending on education, police, and welfare carries important policy implications for a crime-averse politician. This paper's primary interest is in measuring the relative strength of the opportunity cost and government capacity channels. It's secondary contribution is to disaggregate the different channels within government capacity. In the following section, I formalize and combine the existing theories of the alternative pathways.

[1]See Becker [10] for the seminal contribution to the rational actor model of criminal behavior.

[2]*Anomie* refers to the inability of an individual to attain material ends commensurate to her desires which are informed by her social context. Merton [46] proposed that crime could be tied to the frustrations generated by anomie.

[3]See Bursik Jr and Grasmick [18], Sampson and Groves [56], and Beaulieu and Messner [8] for detailed discussions of social disorganization theory. See Wolfgang et al. [64] and Wilson [63] for an overview of the subculture literature.

III. Theoretical Motivation

This section defines the rational citizen and maps the competing theories summarized above into a holistic diagram of influence. Uses this theoretical foundation to summarize the hypotheses it brings to the data.

Defining the Rational Citizen

There are many factors that affect the decision to commit a crime. Crimes can be acts of passion, acts of entertainment, or strategic choices to increase wealth. The deterrence and opportunity cost theories isolate the income-conflict relationship by assuming away competing factors unrelated to the strategic calculations of the rational citizen. I make the same assumption here by assuming that citizens condition their behavior on the expected material rewards associated with the behavior.

In line with Becker [10], consider a representative citizen who is choosing between committing a crime (C) and finding a job (J). Each action has expected returns associated with them which are assumed to be orthogonal.[4] The returns to finding a job are a function of the citizen's education, race, gender, age, and the prevailing local economic conditions in the relevant job market. These conditions include the wage as well as the local government's spending on welfare (w). Similarly, the returns to committing a crime are also a function of the citizen's education, race, gender, and age, as well as the capacity of the local government to capture and punish offenders. I remove the common demographic factors and present a simple utility function for the rational citizen, in which legal income (I) is compared to the returns to predation (π).

$$\mathbb{E}[J] = I + w$$
$$\mathbb{E}[C] = \pi - d$$

The rational citizen therefore conditions their choice of legal versus illegal behavior on the expected returns associated with both. Simple algebra highlights that, all else held constant, legal income I must be greater than or equal to the returns to predation π minus the components of government capacity (deterrence d and redistributive spending w). Setting the combined utility function equal to zero and taking derivatives yields predictions for the expected influence of each factor on the likelihood of committing a crime. Specifically, increases in legal income, deterrence, and welfare all reduce the incentives for the rational citizen to commit a crime while an increase in the returns to predation have a positive effect. At the

[4] Clearly, this independence assumption is too strong. However, for the purposes of this empirical paper, I employ it only for the sake of simplicity and to motivate the hypotheses.

margin, we therefore expect exogenous increases in legal income, deterrence, and redistribution to produce a decline in the crime rate.

Note, however, that the above simplification obscures the important empirical debate at hand: namely does an increase in legal income reduce the crime rate directly or via it's impact on the local government's capacity to provide deterrence and redistributive policies? In the context of the United States, local governments rely on both inter-governmental transfers and local taxes to fund their deterrence and redistributive programs. An increase in legal income expands the tax base, thereby empowering local governments with larger budgets to spend on combating crime. I chart the causal pathways from an exogenous income shock on crime outcomes (c_i) in Figure 1, which summarizes the hypotheses that follow.

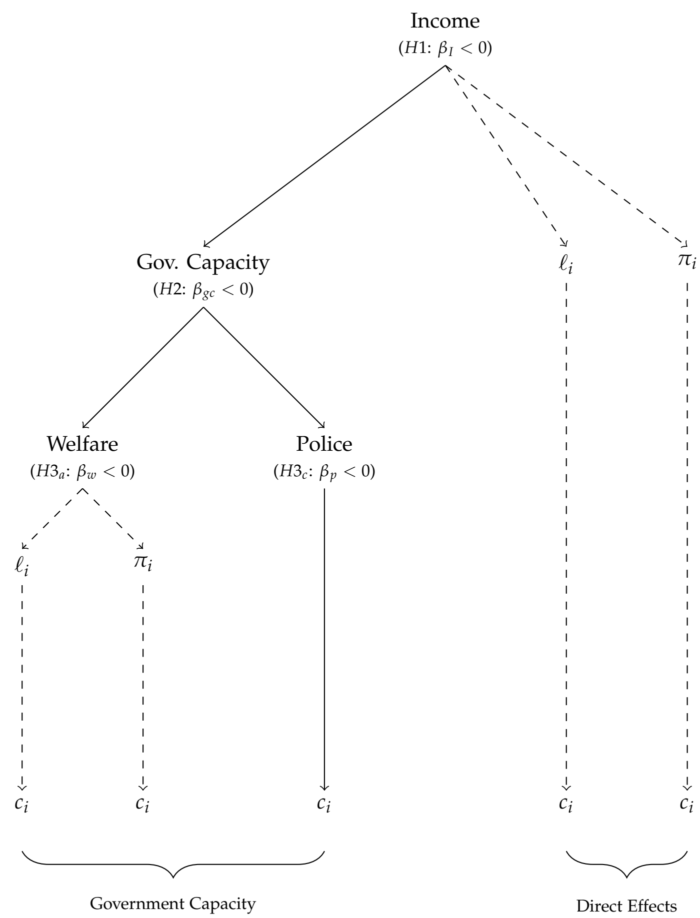

Figure 1: *Mapping the hypotheses. Dashed lines represent the competing theories of opportunity costs and returns to deviant behavior which this paper is unable to disentangle. The index i represents the individual rational actor choosing whether to commit a crime.*

Hypotheses

Existing theories of the direct effect of changes in income on crime present contradictory predictions that hinge on whether the change is relative or absolute. An unequal increase in wealth is predicted to increase crime by presenting more attractive targets for predation while general increases in wealth are theorized to raise the opportunity costs of deviant behavior. I posit that, on average, increases in income reduce crime, regardless of whether the increases correspond to increasing inequality.

Hypothesis 1: An increase in net residential income will cause a decline in crime rates. $\beta_1 < 0$.

I turn to the core empirical test to examine how much of the income-crime relationship is mediated by government capacity. An increase in the government's ability to suppress deviant behavior is predicted to reduce crime since a stronger government is assumed to more effectively combat deviant behavior.[5]

Hypothesis 2: An increase in local government expenditures will cause a decline in crime rates. $\beta_2 < 0$.

Theories from the criminology literature discusses different ways in which government capacity may affect deviant behavior. The primary emphasis is on the ability to deter, with the existing research highlighting police presence and the level of punishment.[6] However, scholars also consider softer forms of government capacity, such as the provision of welfare.[7] Across the board, these influences are theorized to reduce deviant behavior in well-specified regressions.[8]

Hypothesis 3.a: An increase in welfare expenditures will cause a decline in crime rates. $\beta_w < 0$.

Hypothesis 3.b: An increase in police expenditures will cause a decline in crime rates. $\beta_p < 0$.

These hypotheses are summarized in Figure 1. The β symbols represent the coefficient estimating the impact of the variable on crime. The dashed lines represent the theoretical paradox regarding changes in absolute and relative wealth. Absolute changes in income are represented by the opportunity cost channel (ℓ). Relative changes in income are represented by the returns to deviant behavior (π). Without an observable measure of the returns to crime, the dominant effect can only be inferred from the sign on the coefficient.[9] A rigorous analysis of the relative effects within this paradox is left to future research.

[5]See marchese di Beccaria and others [43], Bentham [12], Chadwick [19].

[6]See Ehrlich [28], Viren [62].

[7]See Merton [46], Agnew [1], and Arthur [4].

[8]There is a possible exception for welfare stemming from a debate regarding the conflicting effects of anomie and targets of predation. The strong correlation between welfare and certain demographic characteristics make the variable a proxy for ascribed inequality (see Shaw and McKay [57]). In addition, welfare has been shown to motivate criminal activity as it gets closer to the next payment date (see Foley [31]). As I am primarily interested in the strategic calculations of rational actors, I hypothesize that welfare's effects will be negative.

[9]A net negative effect is theorized to reflect a dominant opportunity cost effect while a net positive effect is theorized to reflect a dominant returns to crime effect.

IV. Methods and Data

This section summarizes the empirical specifications and describes the instrumental variable strategy before turning to a description of the data itself.

Identification Strategy

The motivating theoretical debate is fundamentally over the competing channels of government capacity and the opportunity cost. These pathways can be tested with mediation analysis techniques developed by Baron and Kenny [5]. In this methodological framework, government capacity represents the mediating variable while opportunity costs are represented by the residual pathways. Figure 2 visualizes these channels in a simplified representation of Figure 1.

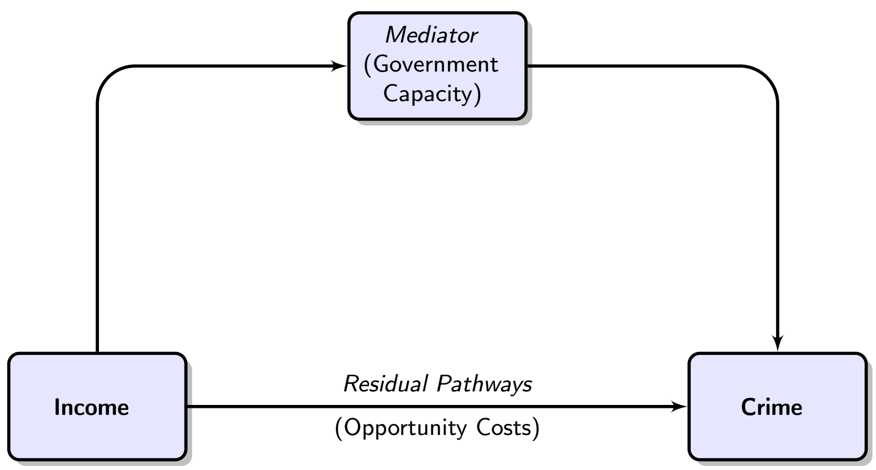

Figure 2: *The mediation framework for conceptualizing the competing theories taken from the conflict literature.*

This section describes the regression models used to test the hypotheses summarized above. It first presents the basic linear model before turning to a brief discussion of the control methods used. The section then describes the instrumental variable strategy before concluding with a description of the meta-regression analysis used to disentangle the competing channels within government capacity.

Basic Linear Model

Equation (1) tests hypothesis 1 which posits that an increase in income should reduce crime. Equation (1) regresses property crime rates on logged net residential income, yielding coefficient estimates that should be interpreted as the response in crime rates to a 1% point increase in income. The basic OLS model is:

$$Crime\ \ Rate_{j,t} = \alpha + \lambda + \beta_I \cdot Income_{j,t} + \delta \cdot \mathbf{X}_{j,t} + \epsilon_{j,t} \tag{1}$$

where the index j, t refers to county j in year t. The error term captures the stochastic component of the model and is likely to be endogenous without further refinement. The α and λ terms represent county and year fixed effects. **X** represents a vector of controls that are motivated by competing theories drawn from the criminology literature. This specification is designed to isolate the strategic calculations undergirding the motivating theoretical debate between contest and opportunity cost theories.

Equation (2) tests hypothesis 2 which predicts that some part of the income-crime relationship is mediated by government capacity. The mediation analysis produces an estimate of the government capacity coefficient that requires two separate calculations. Equation (2.a) presents the regression of government capacity on income. The estimated parameter of interest (ω_I) is then multiplied by the coefficient on government capacity from equation (2) (represented by γ) to estimate the mediation effect. The coefficient on income (β_I) in equation (2) represents the direct effect via the opportunity cost. Determining the relative explanatory power of the competing theories presented by the contest and opportunity cost models requires a comparison of how the β_I coefficient changes from equation (1) to equation (2) after controlling for the mediating effects of government capacity.

$$Gov.\ \ Capacity_{j,t} = \alpha + \lambda + \omega_I \cdot Income_{j,t} + \delta \cdot \mathbf{X}_{j,t}\eta_{j,t} \tag{2.a}$$

$$Crime\ \ Rate_{j,t} = \alpha + \lambda + \beta_I \cdot Income_{j,t} + \gamma \cdot Gov.\ \ Capacity_{j,t} + \epsilon_{j,t} \tag{2}$$

Controls

Unlike in the theoretical discussion presented above, empirical analysis cannot simply assume away competing factors. I add a vector of controls ($\mathbf{X}_{j,t}$) to partial out the effects of the competing criminologic theories of social disorganization, subcultures of violence, and anomie. These competing factors would undermine the applicability of my results to the conflict theories if not controlled for. The controls are drawn from the existing empirical criminology research and include *income maintenance transfers (log), SSI transfers (log), SNAP transfers (log)*, and *federal and state unemployment transfers (log)* to proxy for anomie factors, *% divorced males and females, % college attainment*, and *population density (log)* to proxy for social disorganization factors, and *% African American, % under the age of 24*, and *% single males* to proxy for subculture factors. I also control for unobservable covariates via *county* and *year fixed effects*. In addition, I restrict my analysis to *instrumental* types of crime. Instrumental crimes are those which are perpetuated in pursuit of a material goal and consist of property crimes (motor vehicle theft, burglary, and theft / larceny) as well as robbery. They are theorized to result primarily from strategic calculations.[10]

[10] See Kovandzic et al. [41] for a discussion of the relative applicability of different criminologic theories to different types of crime. *Expressive* crimes are acts of passion typically characterized by violent crimes including murder and assault. These are theorized to be

I contend that the use of controls, fixed effects, and instrumental crimes restrict the incentives facing the US citizen to isolate her purely strategic calculations in the empirical analysis. In essence, my empirical specification reproduces the assumptions of the conflict theories to focus on the same incentives.

Instrumental Variable Analysis

While there have been many attempts to estimate the income-crime relationship in the criminology literature, the vast majority of existing empirical work has been poorly identified. Research has typically relied on cross-sectional data which suffers from spurious correlation and weak causality. Even estimates produced by panel data regressions risk conflating income's effect on crime with crime's effect on income (reverse causality). This paper joins a short list of authors who have exploited exogenous variation in income to capture its causal link to crime.[11]

To account for reverse causality and additional omitted variables, I use a two-stage least squares (2SLS) analysis that exploits exogenous variation in income generated by trade shocks. This strategy is based on recent microeconomics research, specifically work by David et al. [25] who use trade instruments to identify exogenous variation in local labor markets. Following their approach, I use a crosswalk between Standard Industry Classification (SIC) industry codes and Harmonized System (HS) product codes to connect trade flows with industry employment. This connection allows me to estimate import penetration weights (IPW) by county, calculated using a diff-in-diff specification. These import penetration weights represent plausibly exogenous variation in income caused by international trade shocks, including the implementation of free-trade agreements (FTAs) and productivity shocks in foreign trading partners. This instrument satisfies the exclusion restriction under the assumption that FTA implementation and changes in foreign productivity have no effect on crime except through their effect on local labor markets.[12]

There are several ways to calculate the instrument, varying both the industry of interest and the trading partners. Based on first-stage strength, I use changes in manufacturing imports from low-wage countries to predict local changes in income. As presented in Figure 3, there is substantial variation in manufacturing employment shares by county. While the share of employment in manufacturing has been declining since 1982, the heterogeneous nature of this decline contributes power to my instrumental variable analysis. This instrument aligns with the general consensus that increasing international competition in manufacturing has been particularly devastating for US workers.[13] I multiply the instrument by -1 to make facilitate the

unrelated to strategic calculations.

[11] Perhaps the most compelling (and tragically under-cited) example is Bignon et al. [15] who use the spread of the vine-destroying disease Phylloxera throughout France in the 1800s to generate exogenous variation in income. The authors find that income has a negative causal impact on property crime but not violent.

[12] See Appendix B for a detailed description of how these instruments were created.

[13] See Appendix B for a summary of the first stage tests across a variety of related instruments.

interpretation of coefficients as the effects of a 1% increase in income.

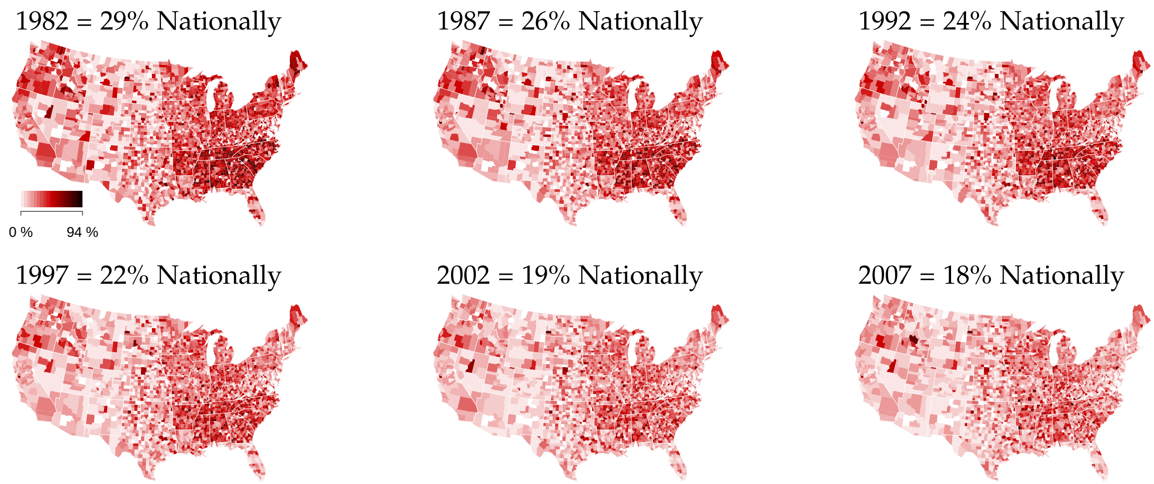

Figure 3: *Manufacturing employment as a share of total employment by county from 1982 to 2007.*

Equation (3.a) is the first-stage regression of income on the trade shock which includes the controls and fixed effects summarized above. Equation (3) is the 2SLS specification where the hat notation indicates predicted values generated by the first stage regression. Note that the 2SLS regressions calculate the relationships of interest using first-differenced measures of the salient variables. This reflects the instrumental strategy which predicts changes in income using changes in import penetration weights.

$$\Delta Income_j = \alpha + \lambda + \pi_1 \cdot \Delta IPW_j + \delta \cdot \Delta \mathbf{X}_j + v_j \tag{3.a}$$

$$\Delta Crime\ \ Rate_j = \alpha + \lambda + \beta_I \cdot \Delta \widehat{Income}_j + \delta \cdot \Delta \mathbf{X}_j + \epsilon_j \tag{3}$$

There is a chance that the exclusion restriction is violated via the effect of trade shocks on government capacity. For example, trade shocks can impact government budgets if they affect the assets in which government funds are invested. If trade shocks affect government capacity independently of income, and government capacity independently affects crime rates (hypothesis 2), then the exclusion restriction will fail unless this channel is partialled out.

While government capacity is included as a control in the mediation analysis, I do not include it in the 2SLS estimates of hypothesis (1). As a result, the estimate of the overall income-crime relationship is potentially biased by the exclusion restriction violation. I do this intentionally in order to generate a baseline estimate against which the mediation results can be compared. As will be shown, the estimate barely changes with the inclusion of government capacity, suggesting that the theorized exclusion restriction violation is negligible.

Meta-Regressions:

I use meta-regression analysis to untangle the various channels through which government capacity mediates the income-crime relationship. First, I match county-year observations based on all available demographic information using Stata's Coarsened Exact Matching (CEM) command[14], creating bins of observations that are similar across observable demographic traits. Within each bin containing 5 or more county-year observations, I regress crime on income using trade shocks as an instrument. These estimates of the income-crime relationship are saved as a vector of coefficients (where n equals the total number of bins). I then regress this vector of coefficients on measures of government expenditure which are also predicted using the trade instrument. These regressions are weighted by the inverse of the standard error of the coefficient estimates.[15] All results are reported as standardized coefficients.

Equation (4.a) presents the regression for calculating the income-crime relationship within each matched bin of county-year observations. Note that the specification is identical to the 2SLS reduced form model except that the county index j has been replaced by bin index b to reflect the estimation of these relationships within matched bins. Equation (4) is the regression of the vector of coefficients on logged measures of government expenditure. This specification includes the controls for the competing criminologic theories summarized above. The hat notation indicates values that have been predicted using the trade instrument.

$$\Delta Crime \;\; Rate_b = \alpha + \beta_I \cdot \Delta \widehat{Income}_b + \epsilon_b \tag{4.a}$$

$$\boldsymbol{\beta_I}_b = \rho_0 + \rho_1 \cdot \widehat{welfare}_b + \rho_2 \cdot \widehat{education}_b + \rho_3 \cdot \widehat{police}_b + \delta \cdot \Delta \mathbf{X}_b + e_b \tag{4}$$

Data

The crime, income, government expenditures, and demographic data was assembled from a variety of portals to the US census, including Censtats, ICPSR, and IPUMS. The trade and industry data used to create the instrument were drawn from the Bureau of Economic Analysis and the County Business Patterns websites. A detailed description of the sources and requisite imputations used to assemble the dataset is presented in Appendix A.

The unit of observation is the county, representing the smallest unit for which detailed data is available. Unfortunately, the criminology literature emphasizes community spill-over effects that may not align with county borders. For example, while a local community in northeastern Vermont may be relatively contiguous with the officially-drawn county lines, the same cannot be said for Hartford County in

[14] This command was developed by Blackwell et al. [16].

[15] This meta-regression specification is inspired by Angrist et al. [2].

Connecticut which contains some of the nation's richest communities along with one of its poorest state capitals. The necessary empirical assumption that county borders describe homogeneous observational units is tenuous at best.[16] Nevertheless, as the smallest observational unit for which repeated observations of the metrics of interest are available, the county represents a substantial improvement over the state or country units typically used in empirical research.

The main results are calculated using net residential income. Existing criminological research has suggested that the rational actor is more responsive to changes in permanent income than temporary (Doyle et al. [27] Dahlberg and Gustavsson [23]). Net residential income is used to represent the permanent component of individual wealth. The robustness section discusses the income-crime relationships calculated using other measures.

Official crime statistics are susceptible to reporting bias since this data is taken directly from the police.[17] Crime victimization surveys overcome many of the measurement issues associated with official statistics. This paper adopts the approach used by Myers [51] that weights official statistics by the propensity to report a crime. These propensity weights are calculated for a variety of demographic categories using a multivariate probit model on the National Crime Victimization Survey data (NCVS). The official county-level statistics are then weighted using the interaction of these demographic-specific weights with the county's own demographic make-up, yielding improved estimates of real crime rates.[18] While the primary results are calculated using property crimes, I include six different types of crime in the robustness checks presented in section VI.

Finally, government capacity is proxied for using a variety of metrics measured at the local level. The basic mediation analysis uses total expenditures while the meta-regressions disaggregate expenditures to capture spending on welfare and police.

V. Results

This section summarizes the results of the empirical analysis. The basic OLS results testing hypothesis 1 are compared to the 2SLS results to determine the extent of bias in poorly identified regressions. The paper then turns to Hypothesis 2 which tests the mediating effects of government capacity on the income-

[16] See McCleary et al. [45], O'Brien [55], and Smith [59] for comprehensive discussions of the unit of analysis issue in criminology empirical work.

[17] Since the propensity to report crimes likely varies systematically with wealth, official statistics are unreliable. Furthermore, police classifications can vary with budgets and policy oversight, leading to certain types of crimes suffering from greater measurement error than others (see Nagin and Paternoster [53]). Often a collection of crimes committed by a single individual at the same time will be recorded as the most serious, leading to a systematic underrepresentation of less severe crimes in the official data (see Becsi [11]). Similarly, victims of assault or petty theft are less likely to report these crimes than victims of motor vehicle theft or burglary for a variety of psychological reasons.

[18] For more information, please refer to online Appendix A.

crime relationship. Finally the meta-regression results are presented, highlighting the channels by which government capacity affects deviant behavior and speaking to hypothesis 3. The results of the OLS, 2SLS, and mediation analyses are ordered by the simple bivariate regression (column 1), the inclusion of controls (column 2), county and year fixed effects (columns 3 and 4), and the fully specified model (column 5).

Hypothesis 1

Both the basic OLS and the 2SLS specifications yield significant negative coefficients in the fully-specified model. However, there is evidence of endogeneity in the OLS model, suggesting that the majority of existing empirical work in the criminology literature is poorly identified.

OLS Results

A simple bivariate OLS regression (presented in Table 1) produces a significant *positive* relationship, as shown in column (1). The inclusion of controls proxying for competing criminologic theories reduces the relationship to a precisely estimated zero as shown in column (2). It is not until the fixed effects are included that the relationship turns significantly negative as shown in columns (3) and (4). The fully-specified regression in column (5), while significant and negative at the 99.9% level of confidence, predicts a reduction in property crimes of only 0.7 offenses per 10,000 people in response to a 1% increase in net residential income.

Table 1: *OLS Regression of Property Crime on Net Residential Income.*

	(1) Bivariate	(2) Controls	(3) w/ County FE	(4) w/ Year FE	(5) Combined
Net Residential	.010***	-.000	-.014***	-.004***	-.007***
Income (log)	(.000)	(.001)	(.001)	(.001)	(.002)
Constant	-.067***	-.066***	.210***	-.096***	.092***
	(.004)	(.007)	(.016)	(.006)	(.019)
Controls	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
County FE	No	No	Yes	No	Yes
N	17980	17980	17980	17980	17980
R^2	.183	.319	.803	.457	.811

Notes: Standard errors in parentheses are clustered at the county level. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

2SLS Results

As depicted in Table 2, estimating the income-crime relationship with the 2SLS model yields a much stronger negative coefficient. The estimate presented in column (5) suggests that a 1% increase in net residential income increases the reduction in property crimes by roughly 6.4 offenses per 10,000 individuals. This result is an order of magnitude larger than the estimate for the simple effect presented in Table 1. The coefficient estimates are consistently negative across the models with the exception of using only county fixed effects (column 3). All first-stage regressions yield F-statistics well above the requisite threshold of 10, as displayed in the top panel of Table 2. The strength of the first-stage regressions limits the concern for bias as discussed by Stock and Yogo [60].

Table 2: *IV Regression of Property Crime Rate on Net Residential Income. Instrumented with manufacturing import penetration weights from low wage countries.*

	(1) Bivariate	(2) Controls	(3) County FE	(4) Year FE	(5) Combined
First Stage: Dependent Variable = Δ Net Residential Income (log)					
IPW IV	7.145***	3.090***	10.730***	2.037***	3.563***
	(.888)	(.616)	(1.140)	(.587)	(.771)
F-Statistic	[64.67]	[25.14]	[88.62]	[12.05]	[21.36]
Reduced Form: Dependent Variable = Δ Property Crime Rate					
IPW IV	-.447***	-.507***	.062	-.677***	-.227*
	(.074)	(.079)	(.094)	(.089)	(.102)
2SLS: Dependent Variable = Δ Property Crime Rate					
Net Residential	-.063***	-.133***	.006	-.332***	-.064*
Income (log)	(.011)	(.039)	(.009)	(.092)	(.031)
Controls	No	Yes	No	No	Yes
County FE	No	No	Yes	No	Yes
Year FE	No	No	No	Yes	Yes
N	14983	14983	14983	14983	14983

Notes: Standard errors in parentheses are clustered at the county level. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

These results confirm hypothesis 1 which predicts that an exogenous increase in income should produce a decline in deviant behavior. Comparing the OLS and 2SLS findings highlights the importance of isolating exogenous variation in income using instrumental variables. The 2SLS results are an order of magnitude larger than the OLS results. The majority of existing empirical work that fails to use well-identified models is therefore called into question.

Hypothesis 2

Theory proposes two alternative channels through which the negative income-crime relationship manifests. The first holds that increases in income improve the legal opportunities for making money, thereby directly increasing the opportunity costs associated with criminal behavior. The competing theory does not challenge the overarching opportunity cost framework. Rather it suggests that increases in income expand the available tax base from which governments draw their revenues. This increased government capacity enables the state to more effectively deter deviant behavior.

Mediation Analysis - General Government Expenditures

Table 3 summarizes the estimated mediation effect of total expenditures on property crime. The average causal mediation effect (ACME, represented by $\bar{\delta}$) estimates are reported below the raw estimates. Calculating the true ACME involves multiplying the coefficient from the regression of total expenditures on income (ω_I) by the coefficient in the main regression (γ). The pure direct effect (or the residual pathways, represented by $\bar{\zeta}$) is simply the coefficient on income produced in the fully-specified mediation specification.

$$Total\ \ Expenditures_j = \alpha + \lambda + \omega_I \cdot Income_j + \delta \cdot \mathbf{X}_j + \epsilon_j$$

$$Property\ \ Crime_j = \alpha + \lambda + \beta_I \cdot Income_j + \gamma \cdot Total\ \ Expenditures_j + \zeta \cdot \mathbf{X}_j + v_j$$

$$\bar{\delta} = \omega_I \cdot \gamma \ \ ; \ \ \bar{\zeta} = \beta_I$$

Column (5) of Table 3 presents the fully-specified results. As shown, the coefficient on net residential income has barely moved although it has lost some of its precision. Meanwhile, the coefficient on total expenditures is a positive 0.003 and is significant at the 95% level of confidence. Employing the calculations summarized above, the ACME estimate is -0.003.

These results demonstrate that the majority of income's negative causal effect on property crime is transmitted directly. What little is mediated by government expenditures is negative and significant but only a fraction of the direct effect. This finding supports hypothesis 2 which predicts that government capacity has a negative effect on crime. However, the comparison of the coefficient sizes suggests that the government capacity channel is negligible compared to the opportunity cost channel. This conclusion is reinforced by the comparison of the direct effect in Table 3 with the 2SLS estimate presented in Table

Table 3: *IV mediation regression of property crime on total expenditures and net residential income.*

	(1) Bivariate	(2) Controls	(3) County FE	(4) Year FE	(5) Combined
(3.a): Dependent Variable = Δ Government Expenditures (log)					
Net Residential Income (log)	2.10*** (.212)	.904+ (.470)	2.491*** (.220)	-2.960** (1.037)	-1.296* (.594)
(3): Dependent Variable = Δ Property Crime Rate					
Net Residential Income (log)	-.101*** (.016)	-.207*** (.050)	-.008 (.011)	-.244*** (.052)	-.063+ (.033)
Total Expenditures (log)	.018*** (.002)	.018*** (.004)	.006*** (.001)	.030*** (.006)	.003* (.001)
ACME ($\omega_I \cdot \gamma$)	.038	.017	.014	-.089	-.003
Controls	No	Yes	No	No	Yes
Year FE	No	No	No	Yes	Yes
County FE	No	No	Yes	No	Yes
N	14983	14983	14983	14983	14983

Notes: Standard errors in parentheses are clustered at the county level. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Both regressions uses 2SLS to instrument net residential income using manufacturing imports from low-wage countries. Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

2. The coefficient has moved from -0.064 to -0.063, suggesting that the intentionally omitted variable of government capacity in Table 2 has no effect on the estimation of the income-crime relationship.

Government Capacity and Income

Contest theories are predicated on the assumption that income and government capacity are positively related. As depicted in the top panel of Table 3, the opposite is shown to be true in the US context. In the fully specified regression, the relationship between income and total expenditures at the county-level are inversely related. The negative ACME is the result of income and government capacity being negatively related and government capacity and crime being positively related.

These results highlight the external validity concerns of using US data to test conflict theories - namely that wealth redistribution ensures that the poorest counties receive transfers from the state and federal levels. However, it seems unlikely that redistribution could fully explain the negative relationship between total expenditures and income. An alternative explanation is that the exclusion restriction for the trade instrument is violated in the regression of government capacity on income. As discussed above, it is possible that trade shocks directly affect government coffers. In this case, the causal estimation of equation (3.a) is violated. To test the robustness of the (3.a) specification to different instruments, I re-run the regression using a Bartik instrument which is less likely to be correlated with government capacity (Bartik [6]).[19]

Table 4 re-estimates the results using the Bartik instrument to generate exogenous variation in income in estimating equation (3.a). As shown, the top-panel yields positive coefficients in line with the assumptions of the contest theories. However, the ACME estimates are also positive, overturning hypothesis (2). Note that the fully-specified regression no longer yields a significant coefficient estimate for equation (3.a). Furthermore, the ACME is now positive and weakly significant.

[19]For a full discussion of the Bartik Instrument, see Appendix B.

Table 4: *Mediation Regression Instrumented with the Bartik Instrument.*

	(1) Bivariate	(2) Controls	(3) County FE	(4) Year FE	(5) Combined
(3.a): Dependent Variable = Δ Government Expenditures (log)					
Net Residential	2.00***	2.28**	3.18**	1.02+	4.58
Income (log)	(.365)	(.737)	(1.02)	(.563)	(4.24)
(3): Dependent Variable = Δ Property Crime Rate					
Net Residential	-.101***	-.207***	-.008	-.244***	-.063+
Income (log)	(.016)	(.050)	(.011)	(.052)	(.033)
Total Expenditures	.018***	.018***	.006***	.030***	.003*
(log)	(.002)	(.004)	(.001)	(.006)	(.001)
ACME ($\omega_I \cdot \gamma$)	.037	.042	.018	.031	.012
Controls	No	Yes	No	No	Yes
Year FE	No	No	No	Yes	Yes
County FE	No	No	Yes	No	Yes
N	14983	14983	14983	14983	14983

Notes: Standard errors in parentheses are clustered at the county level. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Both regressions uses 2SLS to instrument net residential income using manufacturing imports from low-wage countries. Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

These results combine to suggest that, while there is some mediating effect of total expenditures on the income-crime relationship, it is small relative to the opportunity cost effect. Unfortunately, the inability of the trade instrument to satisfy the exclusion restriction for the income-crime and income-government relationships has limited the conclusions that can be drawn from basic mediation analysis. Robustness checks using a Bartik instrument flip the signs on the ACME estimates.

Hypothesis 3

A different approach to testing heterogeneous effects is via meta-regression analysis. This technique overcomes the identification issues associated with the trade instrument by testing whether heterogeneity in types of government expenditure across counties explains heterogeneity in the income-crime relationship.

Overview of the Stratified Results

Stratification on observable covariates produces 282 bins of matched county-years. The number of individual units within each bin ranges from 5 to 118. There is substantial heterogeneity in the income-crime relationships among the bins. Certain bins corroborate the 2SLS results with significant negative slopes while others report significantly positive or null relationships. The meta-regression results attempt to explain this heterogeneity with local government expenditures on police, welfare, and education.

General Proxies for Government Capacity

Table 5 presents the meta-regression results for the income-crime vector of coefficients. This vector of coefficients is regressed on total revenue and total expenditures as well as the proxies for competing theories used in the preceding analyses. Standardized coefficients are presented in brackets.

Table 5: *Meta-Regression Results:*

	(1) Bivariate	(2) General Exp.	(3) Controls	(4) Disagg.	(5) Combined
Total Revenue	-.074*	-.018	.111	-.092	-.006
(log)	(.030)	(.134)	(.133)	(.062)	(.085)
	[-.157]	[-.038]	[.235]	[-.195]	[-.013]
Total Expenditure		-.052	-.163		
(log)		(.120)	(.124)		
		[-.123]	[-.385]		
Police Expenditure				-.054+	-.071+
(log)				(.031)	(.038)
				[-.126]	[-.163]
Welfare Expenditure				.008	.001
(log)				(.012)	(.014)
				[.042]	[.008]
Controls	No	No	Yes	No	Yes
N	282	282	282	282	282
R^2	.025	.026	.056	.044	.067

Notes: Standard errors in parentheses are clustered by matched bin. Beta coefficients presented in brackets. Regressions weighted by the inverse of the standard error of the coefficient estimates. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

The coefficient on total revenue is negative and significant in the bivariate regression model. However, the addition of total expenditures and controls in columns (2) and (3) reduces this effect substantially, making the coefficient indistinguishable from zero and reducing the standardized effect from -0.16 standard deviations to -0.04σ. Columns (4) and (5) reestimate models (2) and (3) but disaggregate total expenditures to examine the particular impacts of police expenditures, welfare expenditures, and educational expenditures. The results suggest that the main channel of mediation is via police expenditures which strengthens the income-crime relationship by roughly 0.16σ. The negative coefficient is marginally significant at the 90% level and is robust to the inclusion of controls.

It bears highlighting that the R^2 value for the fully-specified model is still quite low at 0.07, suggesting that roughly 93% of the variation in the income-crime relationship is left unexplained by government expenditures. The inability of these variables to explain variation in the coefficients reinforces the conclusion that the majority of the income-crime relationship travels directly via the opportunity cost channel.

VI. Robustness of Findings

This section presents placebo tests to confirm the causal identification of the estimated relationships. In addition, I discuss some additional results that speak to the applicability of US data to conflict theories.[20]

Placebo Tests

The main empirical results rest on the assumption that the 2SLS is well-identified. A major motivation for the use of an instrumental variable analysis is to guard against reverse causality. The placebo results presented here rerun the regressions used to test hypothesis 1 and hypothesis 2 using lagged values of the dependent variables to check for reverse causality. Changes in crime rates between period $t-2$ and $t-1$ should not be affected by changes in income between periods $t-1$ and t. If the income-crime relationship persists in the placebo tests, it would undermine the causal strength of this paper's conclusions.

Table 6: *Placebo Tests using fully-specified OLS, 2SLS, and mediation specifications. Dependent variable is lagged.*

	(1) OLS - Placebo	(2) 2SLS - Placebo	(3) Mediation - Placebo
Net Residential Income (log)	-.009***	-.083	-.079
	(.002)	(.067)	(.067)
Total Expenditures (log)			.001
			(.001)
ACME ($\omega_I \cdot \gamma$)			-.002
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
N	14983	11986	11986

Notes: Standard errors in parentheses are clustered at the county level. Controls include anomie theory proxies (income maintenance transfers, supplemental security income transfers, SNAP transfers, and federal and state unemployment transfers), social disorganization theory proxies (% divorced males and females, college attainment, and population density), and subculture theory proxies (% black, % under the age of 24, % percent single males). Significance levels are coded as follows: + = 90% , * = 95% , ** = 99% , *** = 99.9% .

Table 6 summarizes the placebo tests. The models include the full set of controls and fixed effects. The estimates for both the instrumental variable regressions (columns 2 and 3) find no systematic relationship

[20] Additional robustness checks examining the representativeness of the sample are presented in Appendix C.

between the lagged dependent variable and income. However, the same placebo tests run using the OLS specification find a highly significant estimate of roughly the same magnitude as produced using the un-lagged dependent variable. The placebo tests confirm that the instrumental variable analysis does not suffer from reverse causality in the conclusions drawn for both hypothesis 1 and hypothesis 2. Furthermore, the placebo tests indicate that the OLS result is substantially biased by reverse causality.

Crime and Income Heterogeneity

Rerunning the regressions with different measures of crime on different measures of income highlights some additional results that speak to other theories in the criminology literature. In particular, these results indicate that the use of instrumental crimes is effective at isolating the strategic cost-benefit calculations of interest and suggest that permanent measures of income have a stronger effect on crime rates.

2SLS Robustness

The 2SLS results confirm the assumption that evidence of rational actor calculations is limited to instrumental crimes. Figure 4 presents the 2SLS results for a range of income measures over six categories of crimes.[21] The strong causal relationships estimated using property crime rates disappear completely with violent crime rates. Similar trends are evident for the instrumental crimes of burglary and motor vehicle theft. These patterns provide strong support for the theory that opportunity cost calculations regarding criminal behavior are limited to instrumental crimes (Kovandzic et al. [41]).

Turning to the heterogeneous effects by income variable, the only positive coefficients to be found are on the two measures of proprietor's income, despite being imprecisely estimated. Assuming proprietor's income mainly reflects sedentary shops and businesses, this can be taken as suggestive evidence in support of conflict and crime theories which argue that increases in wealth promote deviant behavior by offering more attractive targets.[22]

Finally, it appears that the aggregate measures of income (i.e., total personal, per capita, and net residential) have larger negative relationships with crime than more temporary measures (i.e., wages and wage supplements). This only constitutes suggestive evidence in favor of the theories posited by Doyle et al. [27] who argue that the income-crime relationship is driven mainly by changes in permanent income.

[21] The violent crime category of homicide and the property crime category of larceny-theft are omitted. Homicides were unable to be adjusted using victimization surveys while larceny-theft coverage is missing for 1992 and 1997.

[22] See Besley and Persson [14] and Grossman [36], Grossman and Kim [37] for an overview of the conflict theories. See Land [42], Cohen and Felson [21], and Brush [17] for an overview of the criminology literature.

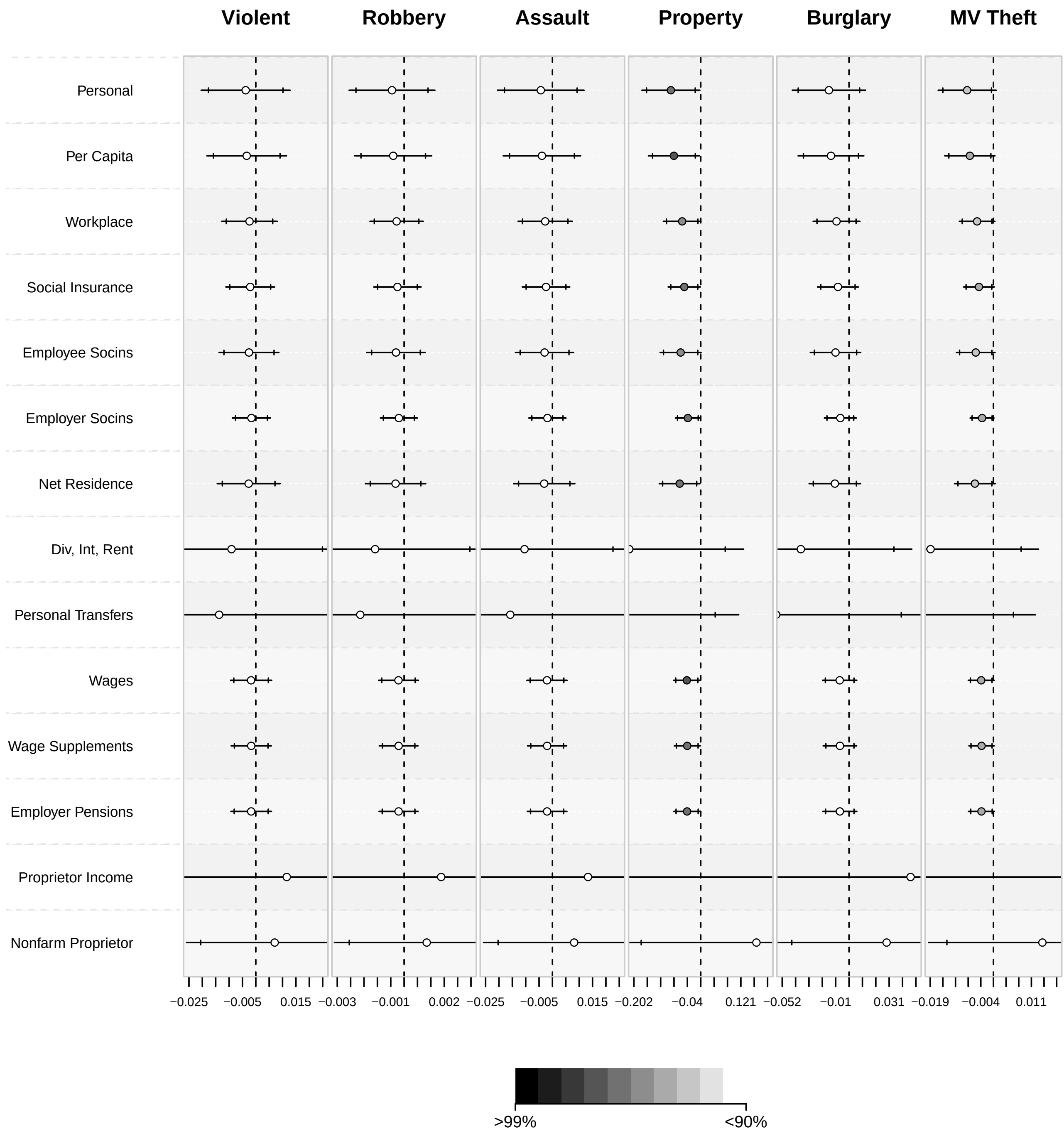

Figure 4: *IV regression of crime rates on income variables. All controls and fixed effects. Standard errors clustered at the county level. Instrumented using import penetration weights of manufacturing imports from low wage countries. 90% confidence intervals represented by horizontal bars. 95% confidence intervals represented by vertical ticks. Significance colored from light to dark gray scale.*

Mediation Robustness

The mediation robustness checks yield similar conclusions (see Figure 5). While the strongest results are found for instrumental crimes, there is also evidence that government capacity has a significant negative effect on robbery which is categorized as a violent (and therefore expressive) crime.

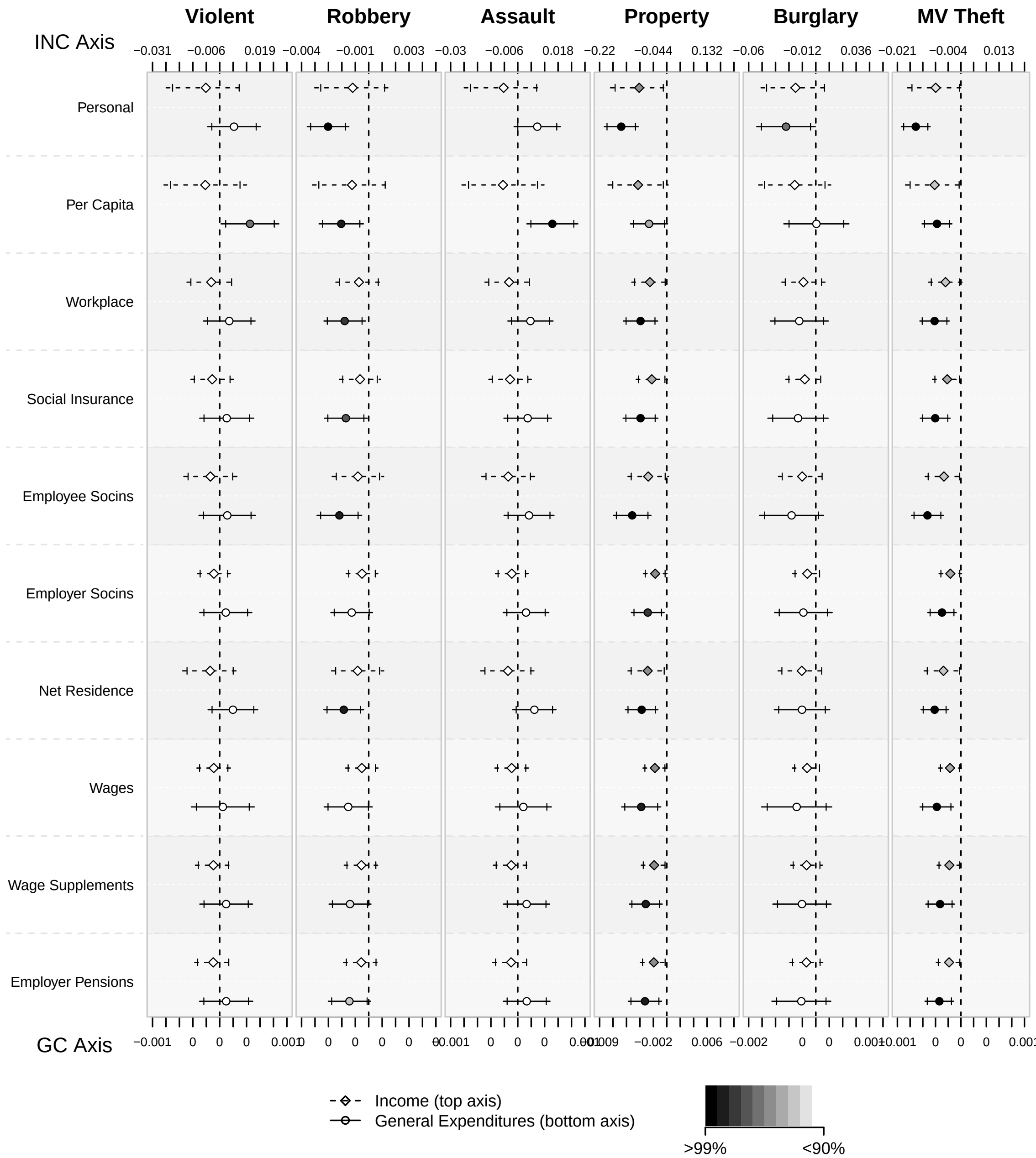

Figure 5: *Mediation IV regression of crime on total expenditure and income variables. All controls and fixed effects. Standard errors clustered at the county level. Instrumented using import penetration weights of manufacturing imports from low wage countries. 90% confidence intervals represented by horizontal bars. 95% confidence intervals represented by vertical ticks. Significance colored from light to dark gray scale.*

Nevertheless, these results suggest that the primary channel by which income affects crime is direct. While government capacity (as proxied for with general expenditures) does have a significant negative effect on instrumental crimes, the magnitude of this effect is negligible compared to the direct effect of income. Furthermore, the inclusion of general expenditures in the 2SLS regression does not dramatically change the sign, size, or significance of the income variable when compared to the basic IV results.

Meta-Regression Robustness

Applying meta-regression techniques to other types of crime yields some interesting additional results. Figure 6 summarizes the results for the basic proxies for government capacity in the top-panel and the disaggregated expenditure variables in the bottom panel. The coefficient estimates are the standardized betas, meaning that the estimates should be interpreted as changes in terms of standard deviations in the income-crime vector of coefficients. Recall that the income-crime relationship itself is weakest in expressive crimes such as violent crimes and assaults. Nevertheless, Figure 6 (following page) highlights the significant impact of police expenditures across all types of crime. These results indicate that the negative impact of a change in income on crime rates is stronger when accompanied by an exogenous increase in police expenditures. Furthermore, the magnitude of this effect is roughly equivalent in terms of standard deviations (about -0.2).

In addition, Figure 6 also indicates that the impact of welfare differs by the type of crime. Welfare expenditures is marginally significant and positive for violent crimes including robbery and assault, implying that the income-crime relationship in these crimes is weaker when accompanied by an exogenous increase in welfare spending at the local level. Conversely, the income-crime relationship is stronger when accompanied by increases in welfare spending for burglary and motor vehicle theft. A more precise estimation of this effect is left to future research.

Robustness Summary

The consistent null results on income in regressions using expressive crimes as the dependent variable supports my assertion that this paper's empirical specification has successfully isolated the strategic incentives discussed in the theory section above. These results indicate that the estimated income-crime coefficients reflect the strategic cost-benefit calculations of interest.

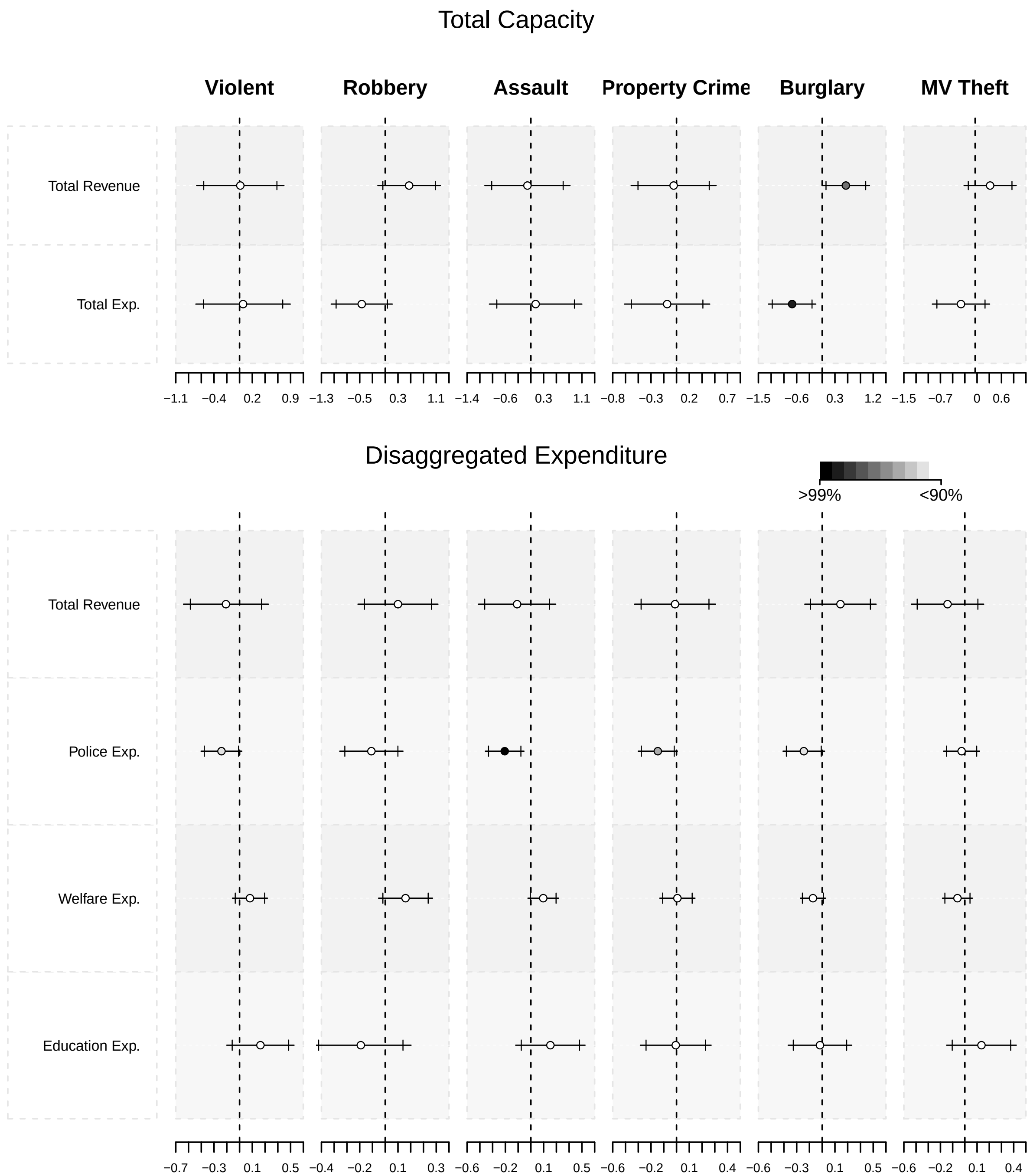

Figure 6: *Meta-regression summary of income-crime relationship calculated using variation in net residential income predicted by changes in manufacturing imports from low-wage countries. All government capacity proxies were predicted using 2SLS regression of capacity variable on net residential income instrumented with manufacturing imports from low-wage countries. Top panel presents results of vector of coefficients regressed on total revenues and expenditures. Bottom panel presents results of vector of coefficients regressed on total revenues as well as disaggregated expenditures. Both panels control for anomie theory proxies, social disorganization proxies, and subculture theory proxies. 90% confidence intervals represented by horizontal bars. 95% confidence intervals represented by vertical ticks. Significance colored from light to dark gray scale.*

VII. Discussion

Having summarized the results, this section discusses their implications for existing criminological theory. This section first discusses the overarching income-crime relationship, ignoring the role of government capacity. It then turns to the mediation results, highlighting the impact of government capacity on the income-crime relationship. Figure 7 maps the complete results onto the theoretical framework.

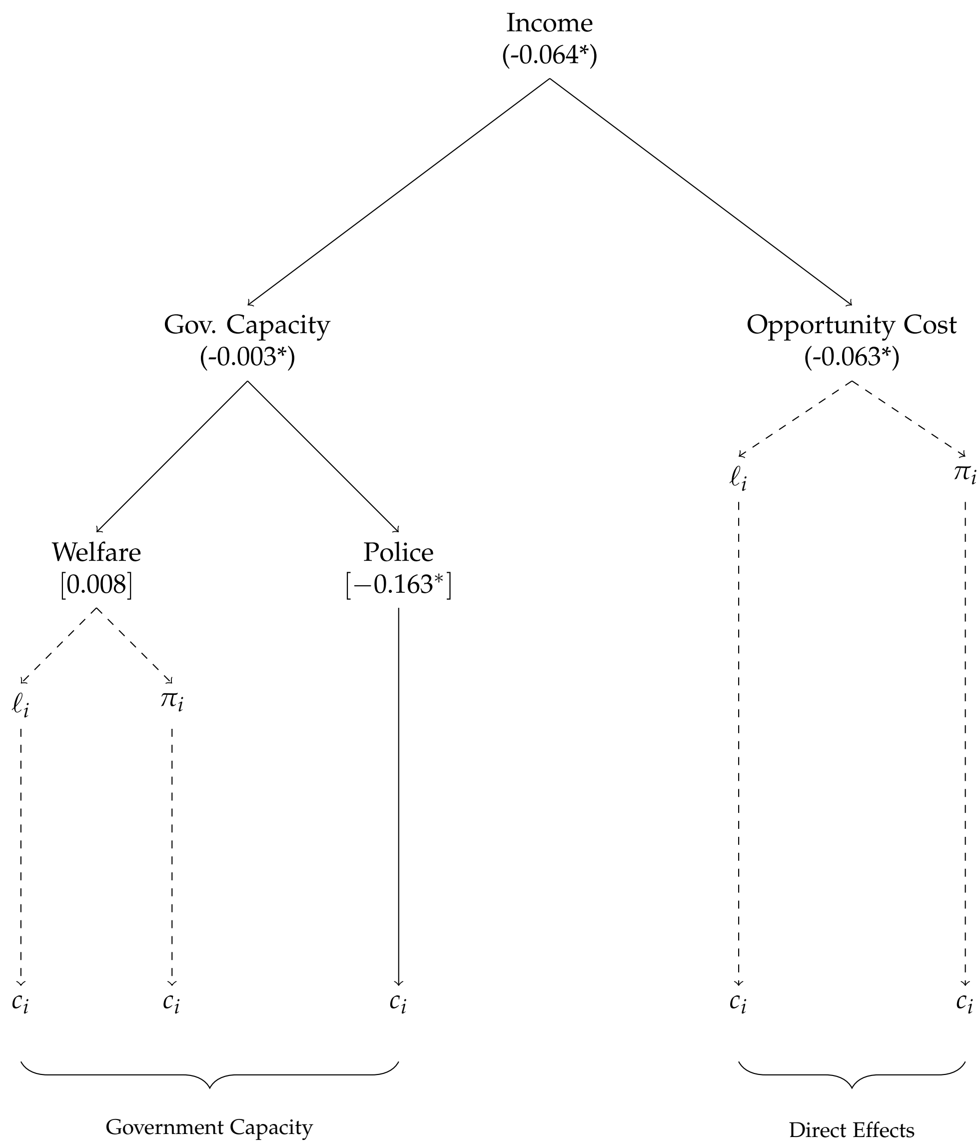

Figure 7: *The combination of instrumental variables analysis, mediation analysis, and meta-regressions is unable to determine the contrasting effects of opportunity costs (ℓ) and returns to criminal activity (π) for either the direct effects or for expenditures on welfare. Bracketed figures are standardized beta coefficients while numbers in parentheses are the coefficients produced by log-level regressions. Asterisks indicate significant results.*

The Income-Crime Relationship:

The findings indicate a negative impact of income on crime rates. This relationship persists after controlling for both competing criminologic theories as well as county and time fixed effects. Under the assumption that the trade instrument's exclusion restriction is not violated, this negative relationship can be interpreted as causal. Specifically, a 1% increase in net residential income is predicted to cause a 0.00064 unit decline in average change in the property crime rate or roughly 6.4 fewer property crimes per 10,000 people.

These results constitute evidence in support of a variety of criminologic theories, all of which find their roots in Becker's 1968 model of the supply of offenses. Whether it be via the direct effect on opportunity costs or the indirect effect via government capacity, income has a causal negative effect on crime rates. This result also aligns with the existing empirical work in the conflict literature which finds a negative causal relationship between GDP and civil war. The coefficient estimate generated by the 2SLS instrumental variable analysis corroborates Hypothesis I - namely that $\beta_I < 0$.

Income, Government Capacity, and Crime:

Adding government expenditures as a control in the 2SLS analysis finds that the direct income-crime relationship remains significantly negative but slightly less so. The coefficient has fallen from roughly 6.4 fewer property crimes per 10,000 people to 6.3. In addition, the mediated coefficient on government capacity is also negative at roughly 0.3 fewer property crimes fewer per 10,000 people for each percentage increase in total expenditures.

Unfortunately, the regression analysis is unable to disentangle the direct effects of opportunity costs and returns to criminal activity. Since the direct effect relationship is negative, I assume that the majority of the income-crime relationship travels via the opportunity cost channel (ℓ). Nevertheless, it is plausible that the well-identified significant coefficient of -0.063 is attenuated by an unobserved positive effect via the returns to crime channel (π). Disaggregating these competing effects is left to future research.

While the mediation analysis confirms Hypothesis II - namely that $\beta_{GC} < 0$, detailed analysis of this effect uncovered unexpected signs in the mediation regression of government capacity on income. Contrary to the contest theories which assume an increase in income causes an increase in state capacity, the US data finds a significant negative relationship. This result may reflect the ability of the US to redistribute funds to local governments most in need. Or it may reflect issues with the trade instrument which may covary directly with government capacity. Substituting the trade instrument with a weaker but more plausibly valid Bartik instrument estimated a positive mediation effect. Nevertheless, the estimate of income's direct effect remained basically unchanged, supporting the conclusion that the majority of the income-crime relationship is explained by the direct effect. Exploring the nuances by which changes in US income affect local government budgets is left to future research.

Within Government Capacity:

Turning to the disaggregation of government capacity proxies, the meta-regression results suggest that the primary mediation effect is via expenditures on police. In addition, there is vague evidence suggesting an additional effect via welfare expenditures for certain types of crime. Given the inability of the data to determine the relative effects via opportunity costs (ℓ) and returns to crime (π), the marginal negative effect of welfare constitutes evidence in favor of the opportunity cost pathway. But without more detailed data on the returns to crime, this interpretation is only suggestive. It is left to future research to determine the competing channels by which welfare affects crime rates.

VIII. Conclusion

Taken together, these results shed light on the ongoing debate over whether the negative causal effect of income on crime rates is primarily a direct effect story (supported by theories emphasizing opportunity cost effects) or a mediation story (supported by theories emphasizing the role of the state). This paper has demonstrated that the majority of the negative income-crime relationship travels directly via opportunity cost channels although it is unable to disentangle the conflicting direct effects of opportunity costs versus returns to crime. Specifically, I conclude that a 1% increase in income causes a decline in property crime of roughly 6.3 offenses per 10,000 people after controlling for the mediating effects of government capacity.

While government capacity does play a role in the income-crime relationship, its effect is modest, contributing only 0.3 fewer offenses per 10,000 people. Nevertheless, this estimate is significant and I find evidence suggesting it is driven largely by spending on police at the local level. The negative effect of net residential income on property crime is enhanced when combined with an increase in police expenditures, reducing the predicted income-crime relationship by 0.16 standard deviations.

Nevertheless, the results do not discount alternative channels entirely. While the effects are small, government capacity does play a role in mediating the income-crime relationship. These significant findings support deterrence theories from the criminology literature.[23] Furthermore, by demonstrating that both channels are active, these findings highlight the need to develop a holistic formal model that incorporates both the direct and mediating channels of the negative impact of income on deviant behavior.

These findings have policy-relevance, particularly given the tools available to politicians. Specifically, I

[23] See Ehrlich [28], Sickles et al. [58], and İmrohoroğlu et al. [50].

find mild but significant support for investing in deterrence capabilities, as evidenced by the beta coefficient on police expenditures. However, the main effect travels via income, suggesting that the primary focus for a politician should be on improving local economic conditions. Interestingly, redistributive programs show no significant effect on criminal behavior, although it bears emphasis that my analysis focused solely on welfare. Alternative types of redistribution, such as training programs that help individuals improve their legal returns, may be more effective.

In addition to contributing to the ongoing debate between deterrence and opportunity cost theories, my results also emphasize the need for a reassessment of the existing empirical research in the criminology literature. The comparison between the OLS and 2SLS results highlight the persistence of endogeneity concerns in robustly-specified regressions using panel data. Much of the existing empirical criminologic work is therefore likely biased by reverse causality and some of the core empirical conclusions should be reassessed using better-identified specifications.

Despite these contributions, my regression techniques are unable to discern how much of the direct effect travels via opportunity costs and how much travels via increased returns to crime. Similarly, I am unable to determine whether the marginally significant effects of welfare expenditures suffer from similar omitted variable bias. I leave these questions to future research.

References

[1] Robert Agnew. Building on the foundation of general strain theory: Specifying the types of strain most likely to lead to crime and delinquency. *Journal of research in crime and delinquency*, 38(4):319–361, 2001. URL `http://jrc.sagepub.com/content/38/4/319.short`.

[2] Joshua D. Angrist, Parag A. Pathak, and Christopher R. Walters. Explaining charter school effectiveness. Technical report, National Bureau of Economic Research, 2011. URL `http://www.nber.org/papers/w17332`.

[3] Peter M. Aronow and Cyrus Samii. Estimating average causal effects under general interference. In *Summer Meeting of the Society for Political Methodology, University of North Carolina, Chapel Hill, July*, pages 19–21. Citeseer, 2012.

[4] John A. Arthur. Socioeconomic predictors of crime in rural georgia. *Criminal Justice Review*, 16(1): 29–41, 1991. URL `http://cjr.sagepub.com/content/16/1/29.short`.

[5] Reuben M. Baron and David A. Kenny. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6):1173, 1986. URL `http://psycnet.apa.org/psycinfo/1987-13085-001`.

[6] Timothy J. Bartik. Who benefits from state and local economic development policies? *Books from Upjohn Press*, 1991. URL `http://ideas.repec.org/b/upj/ubooks/wbsle.html`.

[7] Robert H. Bates. *When things fell apart: state failure in late-century Africa*. Cambridge University Press, 2008. URL `http://books.google.com/books?hl=en&lr=&id=cjeQOriZ_UkC&oi=fnd&pg=PA3&dq=bates+when+things+fell+apart&ots=4QehtDYK-d&sig=HqCOOOcllcz8gwq8r7LMibuFxbc`.

[8] Mark Beaulieu and Steven F. Messner. Assessing changes in the effect of divorce rates on homicide rates across large US cities, 1960-2000: Revisiting the chicago school. *Homicide Studies*, 2009. URL `http://hsx.sagepub.com/content/early/2009/12/15/1088767909353020.short`.

[9] Bernd Beber and Christopher Blattman. The industrial organization of rebellion: The logic of forced labor and child soldiering. *NYU and Yale University*, 2010. URL `http://www.gwu.edu/~igis/assets/docs/Blattman_Paper.pdf`.

[10] Crime Becker. Punishment: An economic approach, 76j. *Pol. Econ*, 169(10.2307):1830482169, 1968.

[11] Zsolt Becsi. Economics and crime in the states. *Economic Review*, (Q1):38–56, 1999. URL `http://ideas.repec.org/a/fip/fedaer/y1999iq1p38-56nv.84no.1.html`.

[12] Jeremy Bentham. *The rationale of punishment*. 1830. URL `http://books.google.com/books?hl=en&lr=&id=7bhCAAAAcAAJ&oi=fnd&pg=PA1&dq=bentham+1830+crime&ots=X1OquhcF4L&sig=2lWn4zM7Iw8uBLfsiUeC4uuA0zM`.

[13] Timothy Besley and Torsten Persson. Wars and state capacity. *Journal of the European Economic Association*, 6(2-3):522–530, 2008. URL `http://onlinelibrary.wiley.com/doi/10.1162/JEEA.2008.6.2-3.522/abstract`.

[14] Timothy Besley and Torsten Persson. The logic of political violence. *The quarterly journal of economics*, page qjr025, 2011. URL `http://qje.oxfordjournals.org/content/early/2011/08/03/qje.qjr025.abstract`.

[15] Vincent Bignon, Eve Caroli, Roberto Galbiati, and others. Stealing to survive: crime and income shocks in 19th century france. 2011. URL `http://halshs.archives-ouvertes.fr/halshs-00623804/`.

[16] Matthew Blackwell, Stefano M. Iacus, Gary King, and Giuseppe Porro. cem: Coarsened exact matching in stata. *Stata Journal*, 9(4):524–546, 2009. URL `http://air.unimi.it/handle/2434/199241`.

[17] Jesse Brush. Does income inequality lead to more crime? a comparison of cross-sectional and time-series analyses of united states counties. *Economics Letters*, 96(2):264–268, 2007. URL http://www.sciencedirect.com/science/article/pii/S0165176507000201.

[18] Robert J. Bursik Jr and Harold G. Grasmick. Economic deprivation and neighborhood crime rates, 1960-1980. *Law & Soc'y Rev.*, 27:263, 1993. URL http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/lwsocrw27&section=26.

[19] Edwin Chadwick. Preventive police. *London Review*, 1(1829):252–308, 1829.

[20] Mitchell B. Chamlin and John K. Cochran. Social altruism and crime*. *Criminology*, 35(2):203–226, 1997. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.1997.tb00875.x/abstract.

[21] Lawrence E. Cohen and Marcus Felson. Social change and crime rate trends: A routine activity approach. *American sociological review*, pages 588–608, 1979. URL http://www.jstor.org/stable/2094589.

[22] Mark Colvin, Francis T. Cullen, and Thomas Vander Ven. Coercion, social support, and crime: An emerging theoretical consensus*. *Criminology*, 40(1):19–42, 2002. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.2002.tb00948.x/full.

[23] Matz Dahlberg and Magnus Gustavsson. Inequality and crime: Separating the effects of permanent and transitory income*. *Oxford Bulletin of Economics and Statistics*, 70(2):129–153, 2008. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0084.2007.00492.x/full.

[24] Ernesto Dal Bó and Pedro Dal Bó. Workers, warriors, and criminals: social conflict in general equilibrium. *Journal of the European Economic Association*, 9(4):646–677, 2011. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1542-4774.2011.01025.x/full.

[25] H. David, David Dorn, and Gordon H. Hanson. The china syndrome: Local labor market effects of import competition in the united states. Technical report, National Bureau of Economic Research, 2012. URL http://www.nber.org/papers/w18054.

[26] Dominique J.-F. De Quervain, Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. The neural basis of altruistic punishment. *Science*, 2004. URL http://psycnet.apa.org/psycinfo/2004-18100-001.

[27] Joanne M. Doyle, Ehsan Ahmed, and Robert N. Horn. The effects of labor markets and income inequality on crime: evidence from panel data. *Southern Economic Journal*, pages 717–738, 1999. URL http://www.jstor.org/stable/1061272.

[28] Isaac Ehrlich. *The deterrent effect of capital punishment: A question of life and death*. National Bureau of Economic Research Cambridge, Mass., USA, 1973. URL http://www.nber.org/papers/w18.

[29] Ernst Fehr and Klaus M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, pages 817–868, 1999. URL http://www.jstor.org/stable/2586885.

[30] Leo Feler and Mine Senses. Trade shocks and the provision of local public goods. 2014. forthcoming.

[31] C. Fritz Foley. Welfare payments and crime. *The Review of Economics and Statistics*, 93(1):97–112, 2011. URL http://www.mitpressjournals.org/doi/abs/10.1162/REST_a_00068.

[32] Michelle R. Garfinkel. Arming as a strategic investment in a cooperative equilibrium. *The American Economic Review*, pages 50–68, 1990. URL http://www.jstor.org/stable/2006733.

[33] Michelle R. Garfinkel and Stergios Skaperdas. Contract or war? on the consequences of a broader view of self-interest in economics. *The American Economist*, pages 5–16, 2000. URL http://www.jstor.org/stable/25604183.

[34] Scott Gates. Recruitment and allegiance the microfoundations of rebellion. *Journal of Conflict Resolution*, 46(1):111–130, 2002. URL `http://jcr.sagepub.com/content/46/1/111.short`.

[35] Scott Gates. Recruiting child soldiers. In *CSCW workshop,"Techniques of Civil War*, 2004.

[36] Herschel I. Grossman. Robin hood and the redistribution of property income. *European Journal of Political Economy*, 11(3):399–410, 1995. URL `http://www.sciencedirect.com/science/article/pii/017626809400062O`.

[37] Herschel I. Grossman and Minseong Kim. Swords or plowshares? a theory of the security of claims to property. *Journal of Political Economy*, pages 1275–1288, 1995. URL `http://www.jstor.org/stable/2138711`.

[38] Trygve Haavelmo. *A study in the theory of economic evolution*. North-Holland Amsterdam, 1954. URL `http://library.wur.nl/WebQuery/clc/1641749`.

[39] Jack Hirshleifer. The analytics of continuing conflict. *Synthese*, 76(2):201–233, 1988. URL `http://link.springer.com/article/10.1007/BF00869589`.

[40] Jack Hirshleifer. Conflict and rent-seeking success functions: Ratio vs. difference models of relative success. *Public choice*, 63(2):101–112, 1989. URL `http://link.springer.com/article/10.1007/BF00153394`.

[41] Tomislav V. Kovandzic, Lynne M. Vieraitis, and Mark R. Yeisley. THE STRUCTURAL COVARIATES OF URBAN HOMICIDE: REASSESSING THE IMPACT OF INCOME INEQUALITY AND POVERTY IN THE POST-REAGAN ERA*. *Criminology*, 36(3):569–600, 1998. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.1998.tb01259.x/abstract`.

[42] Kenneth C. Land. Property crime rate in the united states: a macrodynamic analysis, 1947-1977. *American Journal of Sociology*, 8:90–118, 1980.

[43] Cesare marchese di Beccaria and others. *An Essay on Crimes and Punishments*. Philip H. Nicklin, 1819. URL `http://books.google.com/books?hl=en&lr=&id=FRDtZqosmnEC&oi=fnd&pg=PR11&dq=beccaria+crime+1819&ots=5O7pS6z1R1&sig=6udO7SdX3cNntEBCCrb2cJecg3o`.

[44] Patricia L. McCall, Kenneth C. Land, and Karen F. Parker. An empirical assessment of what we know about structural covariates of homicide rates: a return to a classic 20 years later. *Homicide Studies*, 2010. URL `http://hsx.sagepub.com/content/early/2010/05/17/1088767910371166.abstract`.

[45] Richard McCleary, Barbara C. Nienstedt, and James M. Erven. Uniform crime reports as organizational outcomes: Three time series experiments. *Social Problems*, pages 361–372, 1982. URL `http://www.jstor.org/stable/800025`.

[46] Robert K. Merton. Social structure and anomie. *American sociological review*, 3(5):672–682, 1938. URL `http://www.jstor.org/stable/2084686`.

[47] Steven F. Messner. Poverty, inequality, and the urban homicide rate: Some unexpected findings. *Criminology*, 20(1):103–114, 1982. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.1982.tb00450.x/abstract`.

[48] Steven F. Messner. Regional differences in the economic correlates of the urban homicide rate. *Criminology*, 21(4):477–488, 1983. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.1983.tb00275.x/abstract`.

[49] Steven F. Messner. Merton's "social structure and anomie": The road not taken. *Deviant Behavior*, 9(1):33–53, 1988. URL `http://www.tandfonline.com/doi/abs/10.1080/01639625.1988.9967766`.

[50] Ayşe İmrohoroğlu, Antonio Merlo, and Peter Rupert. On the political economy of income redistribution and crime. *International Economic Review*, 41(1):1–26, 2000. URL `http://onlinelibrary.wiley.com/doi/10.1111/1468-2354.00053/full`.

[51] Ted Myers. Alcohol and violent crime re-examined: Self-reports from two sub-groups of scottish male prisoners. *British journal of addiction*, 77(4):399–413, 1982. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1360-0443.1982.tb02471.x/abstract.

[52] Samuel L. Myers Jr. Crime in urban areas: New evidence and results. *Journal of Urban Economics*, 11(2): 148–158, 1982. URL http://www.sciencedirect.com/science/article/pii/0094119082900250.

[53] Daniel S. Nagin and Raymond Paternoster. Enduring individual differences and rational choice theories of crime. *Law and Society Review*, pages 467–496, 1993. URL http://www.jstor.org/stable/3054102.

[54] Matthew J. Notowidigdo. The incidence of local labor demand shocks. Technical report, National Bureau of Economic Research, 2011. URL http://www.nber.org/papers/w17167.

[55] Robert M. O'Brien. Metropolitan structure and violent crime: Which measure of crime? *American Sociological Review*, pages 434–437, 1983. URL http://www.jstor.org/stable/2095236.

[56] Robert J. Sampson and W. Byron Groves. Community structure and crime: Testing social-disorganization theory. *American journal of sociology*, pages 774–802, 1989. URL http://www.jstor.org/stable/2780858.

[57] Clifford R. Shaw and Henry D. McKay. Juvenile delinquency and urban areas. *Chicago, Ill*, 1942. URL http://books.google.com/books?hl=en&lr=&id=MyKoAgAAQBAJ&oi=fnd&pg=PA106&dq=shaw+mckay+1942+juvenile+delinquency&ots=HGHc15dUDq&sig=0-PoSqbdW7NsLPEkoBCeHg1H2MM.

[58] Robin C. Sickles, Peter Schmidt, and Ann D. Witte. An application of the simultaneous tobit model: A study of the determinants of criminal recidivism. *Journal of Economics and Business*, 31(3):166–171, 1979. URL https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=63057.

[59] Douglas A. Smith. The neighborhood context of police behavior. *Crime and justice*, pages 313–341, 1986. URL http://www.jstor.org/stable/1147431.

[60] James H. Stock and Motohiro Yogo. Testing for weak instruments in linear IV regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, 1, 2005. URL http://books.google.com/books?hl=en&lr=&id=sLXakrLDloEC&oi=fnd&pg=PA80&dq=stock+yogo+first+stage+IV&ots=5gN5e9DKnd&sig=1PL0mB2M6wB5mmp_vz3RwPTRfTE.

[61] William N. Trumbull. Estimations of the economic model of crime using aggregate and individual level data. *Southern Economic Journal*, pages 423–439, 1989. URL http://www.jstor.org/stable/1059220.

[62] Matti Viren. Modelling crime and punishment. *Applied Economics*, 33(14):1869–1879, 2001. URL http://www.tandfonline.com/doi/abs/10.1080/00036840010017677.

[63] William Wilson. *1987. The Truly Disadvantaged*. Chicago: University of Chicago Press, 2012.

[64] Marvin E. Wolfgang, Franco Ferracuti, and Hermann Mannheim. *The subculture of violence: Towards an integrated theory in criminology*, volume 16. Tavistock Publications London, 1967. URL https://www.ncjrs.gov/App/abstractdb/AbstractDBDetails.aspx?id=83013.

A. Data

This paper draws data from a variety of sources to create its panel of income, crime, and government capacity measures, along with an assortment of controls. This section summarizes the sources of these data and describes the techniques used to concord them. The underlying inspiration for the estimation strategy comes from David et al. [25] and the .do files used to estimate industry exposure by county are amended versions of Dorn's replication materials. The Bartik instrumentation strategy is inspired by Feler and Senses [30] as well as Notowidigdo [54].

Census Data:

The majority of the data was drawn from the US census via a variety of portals. Crime, local government spending, and demographic details were downloaded from the US Censtats database.[24]. Coverage included all 3,198 US counties but varied by year. The data on local government spending were recorded every five years from 1977 to 2007 for revenue totals and from 1977 to 2002 for expenditures. Demographics data was recorded every ten years from 1970 to 2010. Crime data was recorded annually from 1980 to 2008 with the exception of larceny-thefts which were not recorded from 1993 to 2002.

In order to create the panel dataset, I used a simple weighted average imputation method to calculate demographic information for 1982, 1987, 1992, 1997, 2002, and 2007. This technique averaged the two decennial Census data bookending the year in question, weighting the contribution of each by the number of years apart they were. For example, the calculation of the 1982 and 1987 demographic information ran as follows:

$$Dem_{82} = \frac{1990-1982}{1990-1980} * Dem_{80} + \frac{1982-1980}{1990-1980} * Dem_{90}$$
$$Dem_{87} = \frac{1990-1987}{1990-1980} * Dem_{80} + \frac{1987-1980}{1990-1980} * Dem_{90}$$

Demographic controls included age, education, race, and marital status by gender. I did not impute the missing larceny-theft data for fear of biasing my results with an imputed dependent variable. For a full list of the variables, please refer to Figure A.1 (following page). The yellow cells indicate missing years for which I used the above imputation technique.

To capture the missing government expenditure variables for 2007, I used Stata's multivariate normal imputation algorithm. I first merged all datasets by county and year. I then predicted the missing expenditure data using all demographics, income, and transfer measures as well as the government revenue measures. The panel was set by county by year. I ran 50 imputations and then averaged the imputed values into a final dataset.

I use county population figures per year to calculate crime rates for both violent and property crime. To adjust the crime rates to account for systematic underreporting, I follow Myers Jr [52]. Using the National Crime Victimization Survey (NCVS) data from 1992 to 2003, I first run a multivariate probit on the binary variable of whether the respondent reported a crime to the police. Dummies representing a range of demographic characteristics (gender, age, educational attainment, income, race, and marital status) are the regressors and I look at reporting by disaggregated measures of violent and property crime. I calculate the marginal effects of each dummy variable for each year and save these as a separate dataset. These values represent the partialled out independent effects of each demographic characteristic on the propensity to report a crime. I then use these values to create the weighted propensity to report crimes for each county. I do this by calculating the average demographic characteristics by county and using these as weights to get the average propensity to report by county by demographic category. I finally average these weighted averages together to get the final propensity to report for each county for each year. The adjusted crime rates are simply the officially reported crime rate divided by the county's corresponding propensity to report for each year.

[24] http://censtats.census.gov/

Figure A.1: *Full list of variables, coverage, and source.*

Source	Variable Type	Sub-Category	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
CENSTATS	Crime	Violent Crimes																																	
http://censtats.census.gov/		Homicides																																	
		Rapes																																	
		Robberies																																	
		Assaults																																	
		Property Crimes																																	
		Burglaries																																	
		Larceny-thefts																																	
		Motor-vehicle Thefts																																	
BEA	Income	Total Personal Income																																	
http://www.bea.gov/regional		Per Capita Income																																	
		Workplace Income																																	
		Social Insurance Income																																	
		...contributed by employee																																	
		...contributed by employer																																	
		Net Residential Income																																	
		Dividends, Interest, and Rent																																	
		Personal Transfers																																	
		Wages																																	
		Wage Supplements																																	
		...pension contributions																																	
		Proprietor Income																																	
		Nonfarm Proprietor Income																																	
CENSTATS	Government Capacity	Total Government Revenue																																	
http://censtats.census.gov/		...intergov revenue																																	
		...intergov revenue from state																																	
		...tax revenue																																	
		......property tax revenue																																	
		General Expenditures																																	
		...Educational Expenditures																																	
		...Welfare Expenditures																																	
		...Health Expenditures																																	
		...Highway Expenditures																																	
		...Police Expenditures																																	
		...Fire Expenditures																																	
		Total Government Debt																																	
		Total Government Employment																																	
		...Total FTE																																	
		...Total Payroll																																	
CBP	Instruments	Bartik Manufacturing																																	
https://www.census.gov/econ/cbp/		Bartik All Industries																																	
		Trade Manufacturing Imports																																	
		Trade MF Imports from Low-Wage																																	
		Trade MF Imports from China																																	
		Trade Exports																																	
CENSTATS	Demographics	Age																																	
http://censtats.census.gov/		Education																																	
		Race																																	
		Marital Status																																	
		Population Density																																	
BEA	Transfers	Total Personal Transfers																																	
http://www.bea.gov/regional		...from Governments																																	
		...Retirement and Disability Insurance																																	
		...Old-Age, Survivors, and Disability Insurance																																	
		...Medical Benefits																																	
		...Medicare Benefits																																	
		...Public Assistance Medical Benefits																																	
		...Income Maintenance																																	
		...Supplemental Security Income																																	
		...Supplemental Nutrition Assitance																																	
		...Other Income Maintenance																																	
		Unemployment Insurance																																	
		State Unemployment Insurance																																	
		Veterans Benefits																																	
		Veterans Pension and Disability																																	
		Education and Training Assistance																																	
		Current Transfer Receipts from Nonprofits																																	

BEA Data:

Income data and transfer data came from the Bureau of Economic Analysis, Regional Income Division.[25] The BEA income data is disaggregated by industry. Unfortunately, the SIC industry codes are only recorded at the 2-digit level. As such, I rely on the County Business Patterns employment data (described below) to get granular 4-digit SIC industry codes to calculate the instrumental variables. Income and transfer data are available for all years and all counties.

CBP Data:

Detailed data on employment and income at the county level by industry code was downloaded from the County Business Patterns website.[26] These data run from 1986 to 2010. Older CBP data for 1980 was downloaded from ICPSR.[27] I used replication .do files provided by David Dorn on his website to concord the counties to 1990 definitions, concord the 1987 SIC industry classification codes with 1997 and 2007 NAICS codes, and impute missing employment values for disaggregated 3 and 4 digit SIC codes. I then merged the industry data by county and year. To calculate the 1982 values, I followed the weighted average imputation described above, using the 1980 and 1986 figures. Please refer to Appendix B for a detailed description of the instruments.

Attrition:

In addition to the imputations described above, changes in official county codes over time yielded the loss of several observations. When these codes would change, I chose to drop the 6 period set of observations entirely to avoid biasing the results with assumptions about how redrawn county borders should be aggregated. For example, FIPS county code 8014 was created in 2001 as an amalgamation of four neighboring counties in Colorado. Similarly, Yellowstone National Park was its own county in the 1990 Census but was incorporated into two neighboring counties for the 2000 Census.

There were also several counties for which data was systematically unavailable. The majority of these were in Alaska which prompted me to drop the state entirely. Virginia also was the source of systematic data issues as it is the only state to include both counties and independent cities within their FIPS county coding system. While some datasets included information on these cities, others did not. As such, I decided to drop all Virginia independent cities from my analysis. These changes amounted to the loss of 1,086 observations, comprising 6% of the total dataset.

In addition, non-systematic missingness prompted me to drop several variables. For example, the income variables of residential income was not recorded for over 6,000 observations. Despite the potential value of this metric as a proxy for permanent income, I chose to drop it and use net residential income in its place. Similarly, many counties did not record the component of proprietor's income that was attributable to farmers. The full list of dropped variables summarized below in Table A.1.

Table A.1: *Systematic Missing Data: Variables that were dropped*

Variable	Missing Observations	Percent of Total
Residential Income	6,142	34%
Farm Proprietor Income	4,478	25%
Military Medical Insurance	4,588	25%
Railroad Retirement and Disability Benefits	2,171	12%

[25] http://www.bea.gov/regional/downloadzip.cfm
[26] https://www.census.gov/econ/cbp/
[27] http://www.icpsr.umich.edu/icpsrweb/ICPSR/

B. Instruments

Motivation

Given the nature of the crime-income relationship, it is highly implausible that statistical analysis can rest on assumptions of either pure or conditional independence. Even after controlling for all observable covariates and using county and time fixed effects, it is unclear in which direction the causal relationship between income and crime runs. The most sensible assumption is to conclude that income causes changes in crime and crime causes changes in income. The endogenous nature of the income-crime relationship means that controls and fixed effects are insufficient to establish a causal relationship.

To speak to the hypothesized causal link from income to crime, it is necessary to isolate a source of exogenous variation in income. Doing so ensures that, not only is the calculated income-crime coefficient unbiased by unobservable covariates, but also that it represents the causal relationship. Figure B.1 diagrams the assumptions underlying the IV approach. Unobservable covariates are represented by the letter A and the $\emptyset$ symbol indicates a null relationship.

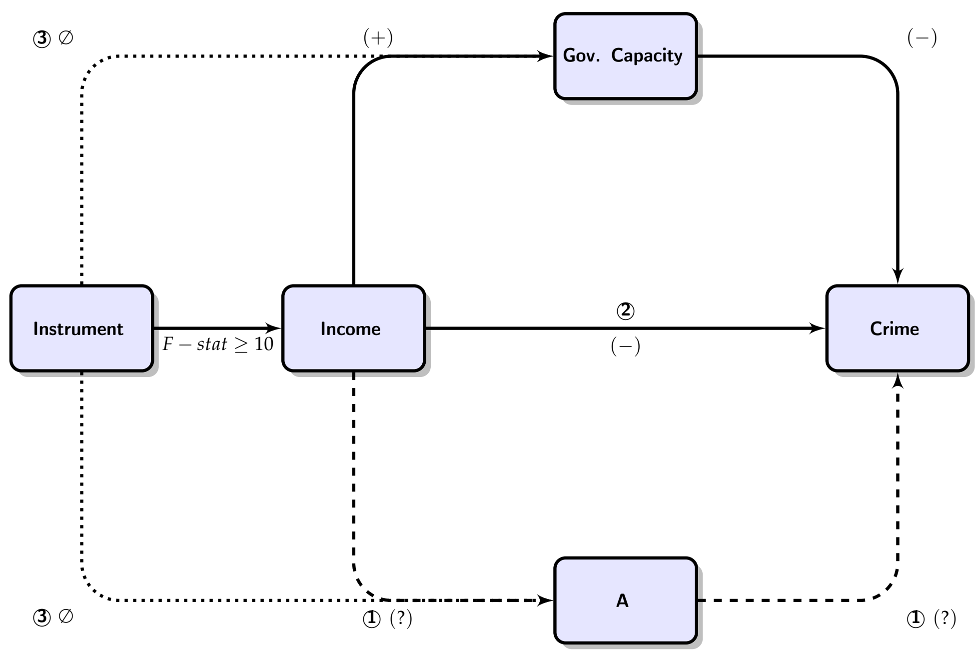

Figure B.1: *The benefits of an IV regression are that ① we don't need to measure A or even understand how it mediates income's effect on crime as long as we assume the instrument cannot possible have an effect on crime except via its effect on income. This benefit buys us ② the ability to claim that the income-crime coefficient is causal. However, these benefits rely on ③ the instrument affecting crime only through its impact on income.*

The crucial assumption of the instrumental variables approach is referred to as the 'exclusion restriction'. As depicted above, it requires assuming ③ - namely that the instrument does not affect crime except through the channel of income. This means that it must not affect income either directly or via unobservable omitted variables. The dotted lines represent relationships that must not exist in order for the instrument to be considered valid and the results of a two-stage least squares regression to be interpreted as causal.

Bignon et al. [15] discuss a number of threats to their identification strategy which are salient to this paper. First, the authors control for the severity of punishment to ensure that the observed increase in crime was the product of the decline in income and not local judges simply becoming harsher. In addition, the authors control for demographic changes to ensure that the disease and attendant decline in agriculture did not lead to a demographic shift that would predict changes in crime. I follow these authors by including a number of controls in my instrumental variable specification that account for competing theories. In addition, I use the number of crimes reported to police as my measurement of crime and adjust it by victimization reports. In so doing, I argue that my data is safe from the severity of sentencing issue that

Bignon et al controlled for.

Inspired by research in the microeconomics literature, this paper identifies two instrumental variables. The first is a set of Bartik Instruments, named after the author of the same name, which isolate exogenous changes in local labor demand using the national employment growth weighted by local employment shares.[28] The second is a set of trade shock instruments that use exogenous variation in import penetration weights caused by FTA implementation and foreign productivity shocks to predict domestic income. The theoretical justification for the trade instruments has been summarized in the body of the paper. The following section describes the Bartik instrument in more detail.

Bartik Instruments:

The Bartik Instrument has been used in much of the urban development and microeconomics literature to identify exogenous variation in labor demand at the local level that is unrelated to local labor supply. The exogeneity of the instrument requires the assumption that county-level changes in the labor market are effected by, but do not effect, national-level changes.

While the Bartik Instrument is well-respected in the urban development literature and microeconomics as a tool to test equilibrium effects following specific shocks, its usefulness in testing the income-crime relationship is shaky. Since the instrument only captures exogenous variation in labor market demand that is unrelated to labor market supply, I must therefore assert that crime is not affected by labor market demand without also being affected by labor market supply. If this assumption does not hold, the Bartik Instrument does not satisfy the exclusion restriction. The Ballentine Diagram in Figure B.2 illustrates the necessary assumption.[29]

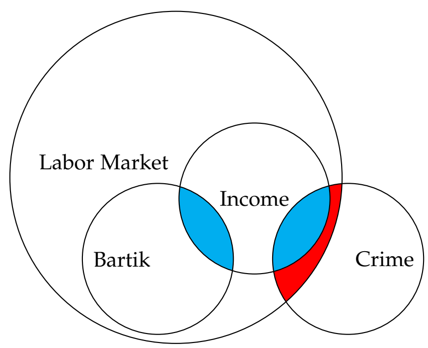

Figure B.2: *The blue shaded area represents the path of exogenous variation in income stemming from changes exclusively in labor demand (labeled "Bartik"). The red shaded area represents the covariance of crime and the overall labor market (including both labor demand and labor supply). As long as this covariance is independent from the Bartik instrument, the exclusion restriction is satisfied.*

However, the exclusion restriction is much more plausible for using the Bartik instrument to estimate the causal relationship between income and government capacity. Changes in local labor demand that do not covary with labor supply are unlikely to vary directly with government capacity except through income. Overall, I include the Bartik instrument only as a robustness check and to speak to the existing urban development literature.

Calculations

Both the trade and Bartik instruments are calculated using first-differenced employment shares by county. For example, a trade shock is measured by the change in imports associated with a certain industry (say, manufacturing) between period $t-1$ and t. The value of this shock affects counties with a larger share of employment in manufacturing, leading to systematically different income levels, also measured as first differences. In essence, both instruments rely on a diff-in-diff framework to generate exogenous variation in income levels.

[28] See Bartik [6]

[29] Thanks to Nathaniel Higgins for introducing me to the Ballentine concept of covariance.

Figure B.3 provides a conceptual diagram of the instruments, using the trade instrument as an example. For the sake of simplicity, imagine import penetration as a binary category of low and high exposure.[30] The timing of the trade shock is also coded [0,1] for a before-after understanding of the approach. In period $t-1$, both the low and high exposure counties have roughly the same average income (allowing for some heterogeneity). After the shock hits, both the low and high exposure counties are affected. However, the magnitude of this effect differs by exposure level. The true effect of the trade shock is calculated as the difference in the high-exposure county's income minus the difference in the low-exposure county's income.

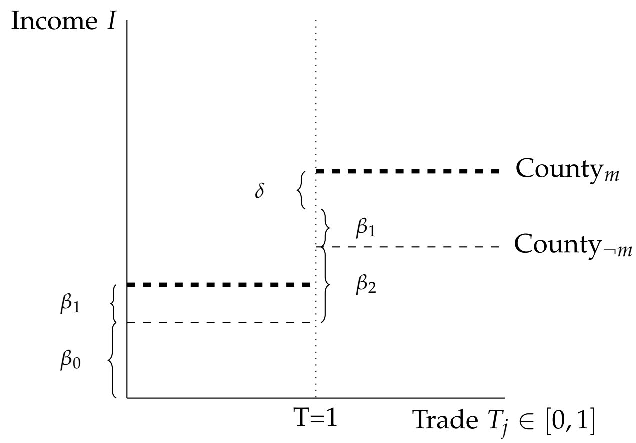

Figure B.3: *Conceptual diagram of diff-in-diff mechanics of trade instruments. For good j corresponding to industry m, trade is at a baseline level in period $T = 0$. The respective incomes of county$_m$ (representing high import exposure or treatment) and county$_{\neg m}$ (representing low import exposure or control) are $\beta_0 + \beta_1$ and β_0 respectively. The effect of the trade shock on the control county is β_2 and the pure effect of the trade shock overall is δ, measured as the difference between treatment and control counties in period $T = 1$ minus the difference in treatment and control counties in period $T = 0$.*

Bartik Calculation:

To calculate the Bartik and trade instruments, I followed the descriptions contained in the primary materials. Bartik's instrument is calculated by interacting county-level employment by industry with changes in nation-wide industrial employment shares. This effectively creates an index of local labor demand which can then be used to predict changes in income. This instrument is widely used in the economics literature which has agreed that it captures exogenous changes in labor demand that are unrelated to local labor supply. As discussed in the body of my paper, my exclusion restriction requirements are more demanding, making the Bartik instrument likely invalid.

To generate the instrument, I first calculate the total US employment by SIC industry codes by year. From this value, I subtract the total county employment by industry and save this as a vector unique values by county by year. Growth rates are then calculated from this vector, giving me the growth rate in US total employment by industry excluding the employment in said industry by county. I then interact this growth rate with the share of county-level employment in each industry by the national growth rate in that same industry. This effectively predicts growth in county labor demand.

Trade Calculation:

The trade instruments require a slightly more detailed description. To understand the theory behind the relationship between trade shocks and income, I borrow notation from David et al. [25]. Let the local labor

[30]This same logic applies to the Bartik instrument for which low-exposure reflects the share of county-level employment shares out of the national total instead of import penetration weights. Note also that this conceptual diagram assumes that a shock prompts an increase in income. In the case of trade shocks, this is actually the reverse of what holds in the data where a decline in tariffs causes a decline in incomes.

demand in region i in industry j (L_{ij}) be a function of the cost per unit produced w_{ij} and the total output Q_{ij}. Total output is itself a function of region i's production that is sold to regions $n \neq i$ and is represented as

$$Q_{ij} = A_{ij} \sum_{n} \frac{X_{nj} \tau_{nij}^{-\theta}}{\Phi_{nj}} \tag{1}$$

where $A_{ij} = T_{ij} w_{ij}^{-\theta}$ represents region i's productivity in industry j as a function of technology (T), production costs (w), and the dispersion of productivity among firms (θ). In addition, τ_{nij} represents the bilateral costs of trade between region n and region i, X_{nj} is the expenditure in destination market n in industry j, and $\Phi_{nj} = \sum_h T_{hj}(w_{hj}\tau_{nhj})^{-\theta}$ represents the level of competition in industry j in region n.

From (1) we can understand how trade shocks affect labor outcomes in region i and industry j. An increase in competition ($\Phi_{nj} \uparrow$) reduces output which lowers labor demand. This increase can be the result of either increased productivity ($T_{hj} \uparrow$ or $w_{hj} \downarrow \Rightarrow A_{hj} \uparrow$) or a reduction in trade costs (i.e., from WTO ascension or signing an FTA, meaning $\tau_{nhj} \downarrow$).

Equation (1) can be operationalized following the strategy of David et al. [25] who sum the output quantities across industries to yield:

$$\hat{Q}_i = -\sum_j \frac{X_{uij}}{X_{uj}} \frac{X_{uhj}(\hat{A}_{hj} - \theta\hat{\tau}_{hj})}{Q_i} \tag{2}$$

This expression is comprised of two components with theoretical value to this paper's empirical strategy. The first expression (X_{uij}/X_{uj}) represents the share of region i in US output u in industry j. An increase in this share increases the exposure of region i to imports from abroad and can be proxied for using the share of regional employment in total national employment in industry j (E_{ij}/E_{uij}).

The second expression represents the magnitude of productivity and trade costs in US imports ($X_{uhj}(\hat{A}_{hj} - \theta\hat{\tau}_{hj})$) relative to the region's total output (Q_i). The larger this expression, the more exposed region i is to import competition from abroad. As above, I proxy for total regional output (Q_i) with total regional employment (E_i). Since it is impossible to obtain detailed information on foreign productivity or expenditures on foreign goods in industry j, I use the change in US imports from abroad in industry j (ΔM_{uhj}). Taking these together and looking over time period t yields:

$$\Delta IPW_{uit} = \sum_j \frac{E_{ijt}}{E_{ujt}} \frac{\Delta M_{uhjt}}{E_{it}} \tag{3}$$

Expression (3) serves as the core instrumental variable for this analysis. It is used to predict changes in income in region i that are exogenously generated by changes in either foreign productivity or trade costs which, in theory, are orthogonal to local crime outcomes. As emphasized by Dorn et al., this measure does not attempt to allocate imports to specific regions. Rather it calculates industry exposure by "[measuring] the *potential* exposure to import competition that local labor markets face due to their industry specialization."[31] The t index represents the measure of the relevant variable at the beginning of the time period while the Δ represents the change in value from the beginning to the end of the same time period.

Trade Instrument Issues

Changes in imports may also reflect changes in domestic demand for foreign goods and services. If domestic demand for imports varies systematically with crime rates, the exclusion restriction fails. Agnew [1] speaks to a related concept in describing how economic goal blockage, by itself, may not constitute a high enough strain to prompt criminal behavior. Specifically, he describes how it is the combination of goal blockage with a perception of injustice that pushes individuals toward extra-legal activities. If increasing imports and the decline in industry-specific incomes they represent are popularly perceived as unjust, this paper's instrumental variable strategy will not control for strain theory's competing predictions.

[31] David et al. [25, pg. 8]

Complete First-Stage Tests

The figures below present the first-stage tests of both the trade and Bartik instruments on all measures of income, all measures of crime, and all measures of government capacity. As depicted, the strongest instruments are the manufacturing imports from low-wage countries and China.

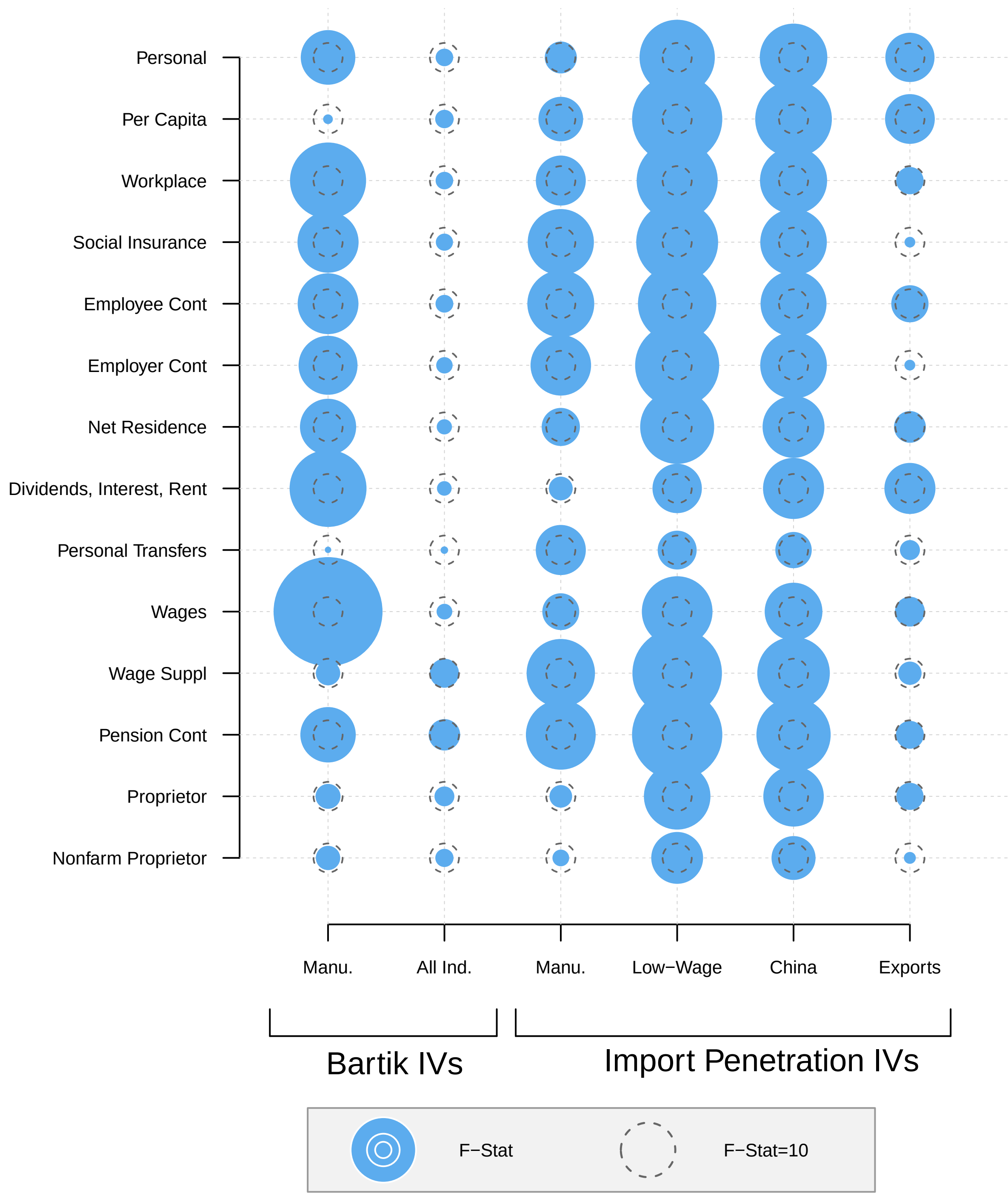

Figure B.4: *First Stage Strength of all instruments on all measures of income. The* ***area*** *of the circle represents the size of the F-statistic.*

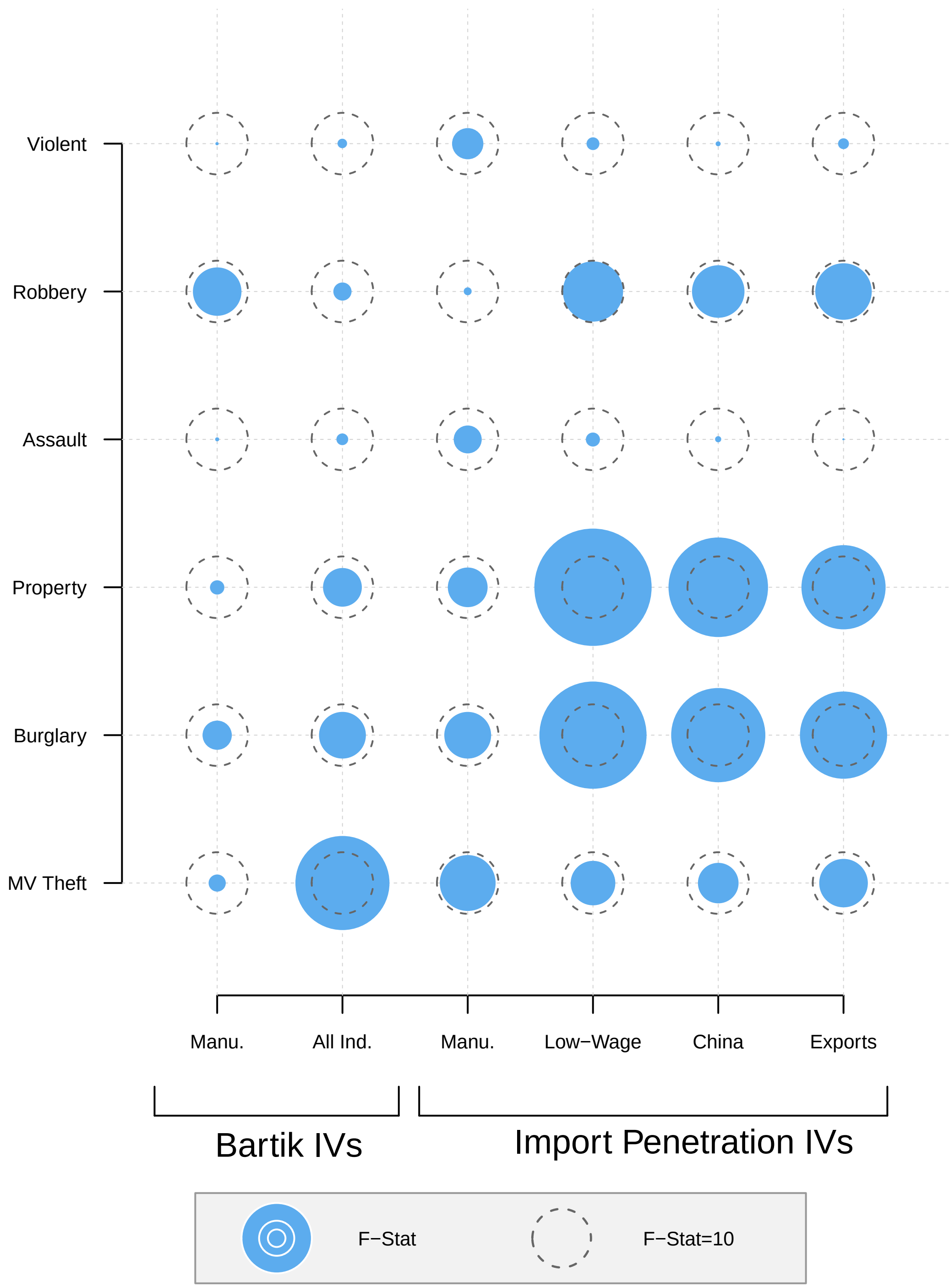

Figure B.5: *First Stage Strength of all instruments on all measures of crime. The* ***area*** *of the circle represents the size of the F-statistic.*

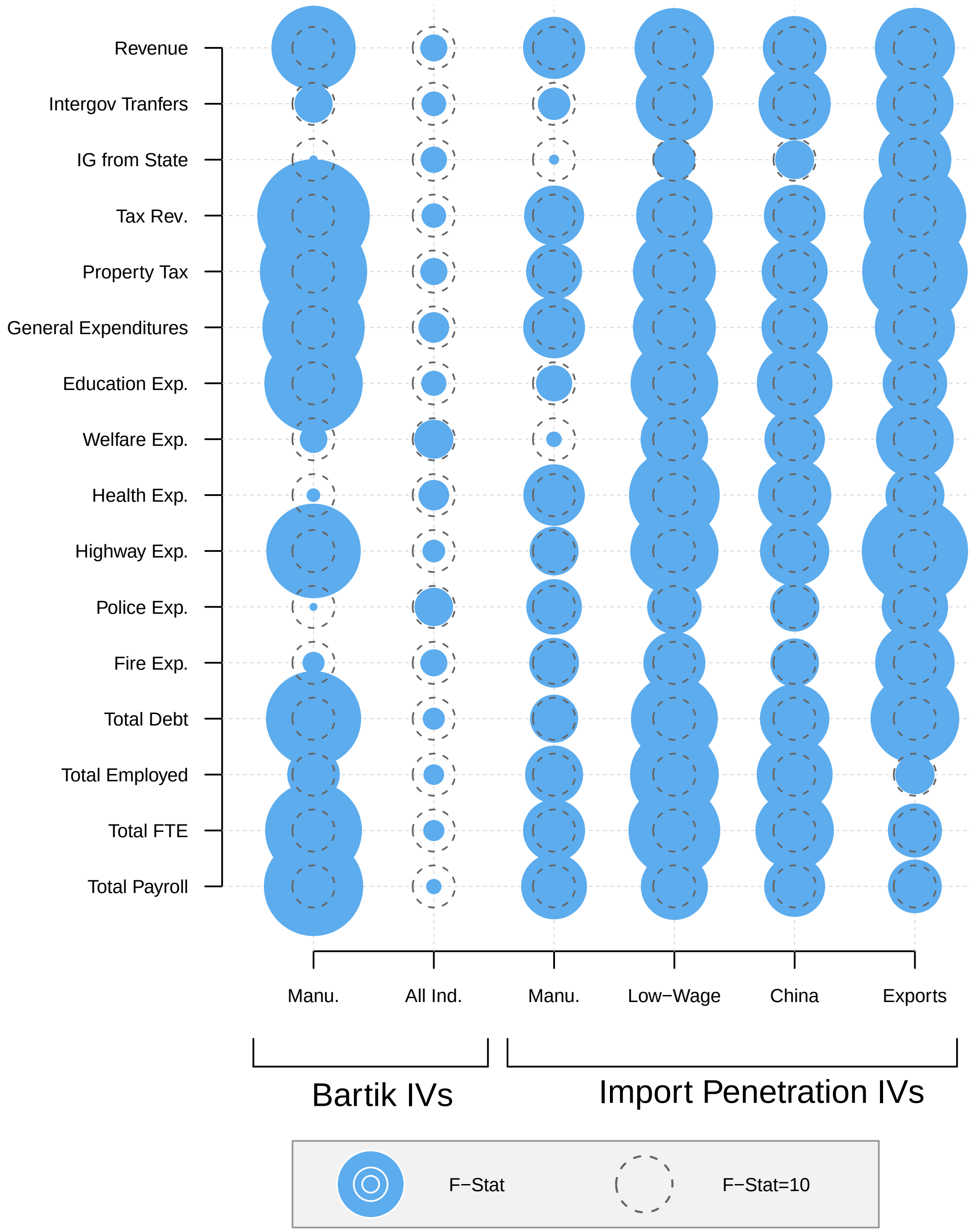

Figure B.6: *First Stage Strength of all instruments on all measures of government capacity. The* ***area*** *of the circle represents the size of the F-statistic.*

C. Representativeness

This paper is predicated on a rational actor framework to model the income-crime relationship with an opportunity cost mechanism. Since salient data is not reported at the individual level, I use counties as the smallest nationally representative units that are observable over time. As such, counties can be understood as proxies for individuals in whom the data generating process is assumed to reside.

It therefore follows that county populations are the most obvious benchmark against which to test the representativeness of my results. In plainer language, if a county's population represents X% of the national total, we would hope that the same county contributes roughly X% to the estimation of the parameters of interest. Figure C.1 presents a heatmap of county population shares of the national total in the dataset. In this figure, darker colors represent larger shares of the national total and are therefore expected to contribute more information to the estimation of the income-crime relationship. However, analyzing the effective contributions by each county to regression estimates highlights some cause for concern.

Take the following reductive example to understand how a representative sample may not aggregate causal effects in a representative manner. Consider an income shock as a binary [0,1] treatment in which 1 represents a positive shock and 0 represents a null or negative shock. For the purposes of this discussion, I will assume away the higher-level endogeneity concerns, allowing us to appeal to the conditional independence assumption (CIA) to estimate the average treatment on treated effect (ATT). Under these assumptions, we have the following:

$$\begin{aligned}\rho_{ATT} &= E_{X|D}\{E[Y_{1i} - Y_{0i}|X_i, D_i = 1]|D_i = 1\}\\ &= E_{X|D}\{\underbrace{E[Y_{1i}|X_i, D_i = 1]}_{\text{observable}} - \underbrace{E[Y_{0i}|X_i, D_i = 1]}_{\text{Counter-factual}}|D_i = 1\}\\ &= E_{X|D}\{E[Y_{1i}|X_i, D_i = 1] - \underbrace{E[Y_{0i}|X_i, D_i = 0]}_{\text{by CIA}}|D_i = 1\}\end{aligned}$$

where D represents the binary [0,1] treatment indicator, Y is the outcome variable, and X_i is a vector of covariates, indexed by individual observations $i \in [1, \ldots, n]$. In principal, we want to calculate the causal effects for each values of X_i and then average these using the distribution of X_i over the treated units. This would yield

$$\begin{aligned}\hat{\rho}_{ATT} &\xrightarrow{a} \sum_x \delta_X Pr[X_i = x|D_i = 1]\\ &= \frac{\sum_x \delta_X Pr[D_i = 1|X_i = x]Pr[X_i = x]}{\sum_x Pr[D_i = 1|X_i = x]Pr[X_i = x]}\end{aligned}$$

where $\delta_X = E[Y_{1i}|X_i = x, D_i = 1] - E[Y_{0i}|X_i = x, D_i = 0]$. In the common tongue, all this describes is a process in which, had we the resources, we could match all observations on observable X and calculate the causal effect of treatment by subtracting the average of the outcome for the untreated observation from the treated average. This vector of causal effects is aggregated via population weighting to arrive at the sample estimate of the parameter of interest.

Unfortunately, (or fortunately if our data is well-behaved), OLS regressions that partial out X to satisfy CIA yield a slightly different aggregation function. Following the notation of Aronow and Samii [3], the parameter estimate produced using regression analysis is

$$\hat{\delta}_R \xrightarrow{a} \frac{\sum_x \delta_X [Pr[D_i = 1|X_i = x](1 - Pr[D_i = 1|X_i = x])]Pr[X_i = x]}{\sum_x Pr[D_i = 1|X_i = x](1 - Pr[D_i = 1|X_i = x])]Pr[X_i = x]}$$

The crucial difference between these two parameter estimates is that the first is weighted by the population while the second is weighted by the variance of X_i with respect to treatment. The variance weighting technique is what is produced in regression output and is, in theory, more efficient than the population weighted estimate. Indeed, if the causal effects were constant over values of X_i, the variance weighting would yield unbiased and precise estimates. However, insofar as the causal effects are heterogeneous over values of X_i, the variance weighting technique will be biased. This is because it privileges values of X_i for which the δ_x estimates are precise.

Figure C.1 depicts the effective sample calculated using the residuals from the fully specified model (including county and time fixed effects, first differenced variables, and a battery of controls). The darker shades indicate greater county weight in the effective sample. As depicted in Figure C.1, there is significant deviation from the population weighted heatmap and the effective sample used to predict causal effects. (Note that the empty white counties represent observations for which key variables were not present and were therefore dropped.)

Figure C.1: *Population versus variance weighting in the full sample (first row) and with the removal of the 15 most biased counties (second row). As depicted, there is substantial deviation from the population weighted representativeness when measuring with variance weights. As such, results should be interpreted with caution.*

For the purposes of representativeness, this paper estimates all regressions using both the full sample as well as a subset in which the 15 most biased counties are dropped. The heatmaps presented in Figure C.1 compare the effective samples with and without these 15 counties. Clearly, their omission yields a more representative effective sample although one with substantial differences from the population weighted nominal sample. The paper presents the results calculated using the less biased sub-sample. However, the results produced by regressions run on the full dataset are similar and available upon request.