The Distinct Roles of Poverty and Higher Earnings in Motivating Crime

Benjamin Ferri and Lia Yin*

September 12, 2022

Click here for most recent version.

Abstract

Does inequality lead to more crime? We develop a new model that articulates how Poverty (the lower tail of the earnings distribution) and Earnings (the upper tail) enter into equilibrium crime rates. In our model, individuals in Poverty have less to lose in the context of criminal punishment, so are less averse to committing crimes in general. The presence of high Earnings (therefore things worth stealing) heightens the expected gain to offenders per crime - but specifically in terms of financial gain, not emotional gain. We estimate our model on a comprehensive panel of U.S. Commuting Zones (1980-2016), deploying novel Shift-Share instruments to correct for reverse causality (of crime on the earnings distribution). Corroborating our hypothesis, we find that high Earnings plays a much larger role in driving crimes that yield financial gain to the offender (various forms of theft) than it does for crimes of emotional gain; while Poverty is a driving force equally across both types of crime. In each case, not accounting for reverse causality would underestimate both effects, often by more than double.

Keywords: crime, poverty, punishment, earnings inequality, shift-share

JEL Classifications: D31, I32, J17, K42

*Ferri: Boston College (email: benjferri@gmail.com), Yin: Boston College (email: yinlb@bc.edu). We are grateful to our advisors Christopher Baum, Donald Cox, Hanno Foerster, and Arthur Lewbel.

1 Introduction

Do the poor have more reason to engage in crime than the rich? And does the presence of higher earnings inspire more crime, by presenting potential offenders with more fruitful opportunities? We model and estimate these distinct forces in a novelly succinct and comprehensive way. As such, this paper articulates the main roles of legitimate (non-crime) earnings in influencing potential criminal offenders' cost-benefit analysis.

Because Poverty[1] operates by making the poor relatively less averse to criminal punishment, it motivates higher rates of crime across the board - both those that are motivated primarily by financial gain to the offender (Financial-Gain crimes),[2] and those motivated primarily by emotional gain (Emotional-Gain crimes).[3] High Earnings, on the other hand, operate by presenting potential offenders[4] with more fruitful opportunities for crime - but fruitful specifically in the sense of financial gain, rather than of emotional gain. These ideas both motivate our simple structural model, and are corroborated in parameter estimates.

The model has two agents - the criminal offender, and the government. The offender's utility weighs his expected gain from crime against his expected loss from associated punishment. The offender's loss coincides with legitimate income that he forgoes when subject to punishment. The government's loss function weighs social losses from punishment against social losses from crime: due to offenders' response, a higher level of expected punishment per crime leads to a lower level of crime. In equilibrium, the criminal offender maximizes his utility and the government minimizes its loss. The equilibrium crime rate in each Commuting Zone[5] is a function of its average inverse earnings (which we call Poverty), average

[1]Poverty (proper noun) is a technical term in our model. Specifically, we use the term Poverty to mean average *inverse* earnings, which is a mathematical object that arises in our structural model. In juxtaposition to Poverty, by high or higher Earnings we mean plain average earnings. Although a plain average, average earnings represent the upper tail of the earnings distribution in that earnings are organically right-skewed.

[2]Becker (1968) referred to "monetary" gain versus "psychic" gain. We make this distinction more explicit, and also alter the word choice slightly for modern usage.

[3]Financial-Gain crimes are Robbery, Burglary, Larceny, and Motor Vehicle Theft. Emotional-Gain crimes are Murder, Rape, and Aggravated Assault.

[4]Offenders may be of any income level.

[5]Commuting Zones are designed to capture local labor markets - see Tolbert and Sizer (1996) and Dorn (2009). We assume they also capture local crime markets.

earnings (Earnings), and controls that include population density and racial demographics.

Criminal punishment plays an essential conceptual role in our model. Although punishment inevitably must play a part in deterring crime, it may be especially relevant in the US context. As shown in Table 1, the US has a much higher incarceration rate than any other industrialized country. Many jurisdictions have mandatory sentencing, three strikes laws, capital punishment, and other legislation that leads to high levels of punishment per crime. All of the above implicitly factor in to the government's loss.

Intuitively, earnings influence crime; but the reverse is also true. Where Poverty or Earnings increase, crime increases in response. However, crime may motivate out-migration of people with higher Earnings, or societal trauma that results in more Poverty. Such cases of reverse causality, if unchecked, may undermine the accuracy and precision of estimates. To get around this, we use Shift-Share (Bartik) instruments, for each of Poverty and Earnings. Shift-Share instruments are constructed by interacting non-local (national) averages for each national occupational category with corresponding historical proportions (shares) of each category at the local (Commuting Zone) level. The resulting instruments mitigate reverse causality bias, because local crime rates cannot affect national averages (of earnings or inverse earnings); nor are local crime rates likely to affect historical occupational shares - even at the local (Commuting Zone) level.

We find that all of our main parameter estimates would be biased downward in OLS. Figure 1 provides a simple causal diagram that helps to explain directions of bias. Poverty motivates crime, and crime can exacerbate Poverty. In other words, the reverse causality has the same sign as the forward causality. Estimated effects of Poverty on crime should be biased downward in OLS in this case. For Earnings, it is the opposite. High Earnings motivate crime, and crime can *decrease* Earnings: OLS may be biased in either direction in this case. Empirically, we find this bias turns out to be downward as well.

One limitation of most literature on the impacts of earnings on crime is that it does not distinguish between the upper and lower tails of the earnings distribution, which may influence crime in different ways. We contribute by modeling both tails of the distribution

(Poverty and high Earnings), and their distinct impacts. A second limitation of much past research is the level of jurisdiction covered. The majority of existing literature fails to account for spillover of criminal activity from neighboring areas, or else retreats to the level of states, which are typically too large to represent local markets. We use Commuting Zones as the units of analysis, thereby retaining the level of granularity that is desirable, while also capturing relevant local spillover.

A third limitation is that this literature rarely has clear causal identification. Measured effects on crime are likely to be biased due to reverse causality.[6] We ameliorate this problem by using Shift-Share instruments for each measure of the earnings distribution at the Commuting Zone level. To construct the instruments, we compute earnings distribution summaries for 14 occupational categories at the national level, and take weighted averages of these at the Commuting Zone level - weighted by the categories' shares of the local economy in a lagged time period. The identifying assumption is that local crime rates do not impact the lagged occupational makeup of the economy in each Commuting Zone - nor the national earnings distribution within each occupation - but only the local earnings distribution within each. Thus we isolate the effects of the earnings distribution on crime, sans the effects of crime on earnings.

Our parameter estimates indicate that the elasticities of Financial-Gain and Emotional-Gain crime are about the same with respect to Poverty: 0.75 and 0.79, respectively. The more in Poverty a person is, the less legitimate earnings he stands to forego via criminal punishment. This effect applies similarly across both types of crime. The elasticities with respect to high Earnings, on the other hand, are far more disperse: 0.93 and 0.39, respectively. The more Earnings there are present in a Commuting Zone, the larger is the expected benefit to offenders per Financial-Gain crime. The impact of Earnings on Emotional-Gain crime is smaller, and not statistically significant in some specifications. That is, expected benefits per Emotional-Gain crime are not as closely associated with the presence of higher Earnings, because such crimes are not financially motivated.

[6] Gould, Weinberg and Mustard (2002) and Enamorado et al. (2016) use Bartik-like instruments, similarly as we do.

The remainder of this paper is organized as follows. Section 2 covers relevant literature. Section 3 describes our data from the Census, American Community Survey, and FBI Uniform Crime Reports. Section 4 explains how we scale crime rates, which is vital for making comparisons across different types of crimes. Section 5 develops our model. Section 6 explains our causal identification in detail. Section 7 discusses descriptive statistics to help familiarize the reader with the data. Section 8 presents our results. Section 9 simulates counterfactual changes in the government's overall level of toughness against crime, exploring the tradeoff that it faces between crime and punishment. Section 10 discusses implications, and section 11 concludes.

2 Related Literature

This paper contributes to the crime literature by highlighting the distinction between Financial-Gain and Emotional-Gain crimes, and most importantly, how lower and higher earnings interact with these two types of crime differently. However, we build on a vast literature of both theoretical and empirical explorations of the economics of crime.

The seminal work of Becker (1968) considers crime in a framework of costs and benefits for the offender, the victim, and the social planner. The offender takes into account the probability of conviction, the punishment if convicted, and other factors when they determine how many offenses to commit within a given time frame. Ehrlich (1970, 1973, 1975) expand upon Becker (1968) and explore the roles that earnings might play in the offender's decision to commit a crime. Witte (1980) builds upon Ehrlich's work and introduces income from illegal activities as a critical factor. Burdett, Lagos and Wright (2003) and Engelhardt, Rocheteau and Rupert (2008) emphasize that crime and income can influence each other and model them as a joint process.

There is a large empirical literature as well. Kelly (2000), Fajnzylber, Lederman and Loayza (2002), Stolzenberg, Eitle and D'alessio (2006), Brush (2007), Choe (2008), Kang (2016), Enamorado et al. (2016) focus on the relationship between inequality and crime.

Measures of inequality include the Gini coefficient, the ratio of mean to median household income, and ratio of 90th percentile to 10th percentile. Levels of jurisdiction include country, state, county, and city. Hipp (2007) is noteworthy in examining the role of racial heterogeneity, and also in attempting to distinguish poverty from inequality per se. Closely related to the Financial-Gain mechanism in our paper, Demombynes and Özler (2005) show that wealthier areas see higher Burglary rates.

Most of the papers we have surveyed support a positive relationship between inequality and crime (Fajnzylber, Lederman and Loayza, 2002, Hipp, 2007, Choe, 2008, Kang, 2016, Enamorado et al., 2016). Notably, (Enamorado et al., 2016) estimate a 36% increase in drug-related homicide rate for each one-point increment in the Gini coefficient. However, there is considerable variation in findings. Some authors have found a positive relationship between inequality crime only for certain types of crime, or only under certain conditions (Kelly, 2000, Brush, 2007, Demombynes and Özler, 2005), or occasionally no relationship at all (Stolzenberg, Eitle and D'alessio, 2006). Our approach may help yield additional insight into this important, yet unsettled topic.

3 Data

This study requires a large dataset. Commuting Zone (CZ) level Poverty and Earnings may be either positively or negatively correlated.[7] It follows that large cross sections may be needed to estimate their independent effects. Moreover, each observation in the cross section is a geographical cell - and therefore a large number of individual person observations are needed, separately within every cell, in order to accurately calculate average earnings and other measures within each. This is a challenge because, in most datasets, geographical location is redacted up to a high level of aggregation (such as state, rather than county). And few datasets would have enough observations to occupy a large number of geographical

[7]A positive correlation would be if CZs in which the rich are very rich also tend to be CZs in which the poor are very poor. Alternatively, for a negative correlation, CZs in which the rich are very rich may tend to be CZs in which the poor are only moderately poor in national terms.

cells, with a large number of observations within each.[8] Finally, in order to constitute a panel, the data must do all of the above for multiple time periods, and in reference to geographical cells that are fixed over time.

IPUMS USA,[9] sourced from the Decennial Census and American Community Survey (ACS), is the largest publicly available compilation of microsamples (samples in which the observations are of individual persons) for the US. IPUMS provides 1% samples of the entire US population. That is, the total number of observations in the sample per time period is about 1% of the national population, so about 3 million (or 2 million, for earlier periods). Observations of individual persons are necessary for flexibility in variable creation. However, we also use Census and ACS summaries at the county level, in addition to IPUMS. In general, IPUMS provides more specific information (variables), while the county level summaries provide more complete coverage. We combine them to yield a panel dataset with a combination of specificity and coverage that neither has on its own.

To constitute locations as referenced in the model, it is necessary to use geographical units that meaningfully capture local labor and crime markets. Although Metropolitan Statistical Areas (MSAs) are the most traditional geographical unit for this purpose, Commuting Zones (CZs) have three major advantages.[10] First, CZs are more comparable over time from 1950 to the present. This is because CZs are defined as groups of counties, and county boundaries have remained mostly fixed, while MSA boundaries have been repeatedly adjusted. Second, CZs cover the entire United States, no matter how remote, while MSAs cover only urban areas. Third, CZs are delineated based on actual commuting patterns (albeit in 1990),[11] which means they are ideal for representing actual local markets. We use the set of 722 CZs that are coterminous with the entire continental US (48 states + DC). Each CZ is 4-5 counties on average, so the 722 CZs cover about 3,200 counties in total.

We obtain county level crime data from the FBI's Uniform Crime Report (UCR) detailed arrest and offense reports (Federal Bureau of Investigation, 1980, 1990, 2000, 2010, 2016).

[8]See Molloy, Smith and Wozniak (2011).
[9]Ruggles et al. (2020).
[10]See Dorn (2009) for more discussion.
[11]See Tolbert and Sizer (1996).

Because Commuting Zones (CZs) are defined as particular clusters of counties, county level data aggregates easily into CZ level data. Data from multiple decades enable us to employ a rich set of regional fixed effects to absorb unobserved information that might be correlated with our variables of interest.

We group index crimes into Financial-Gain and Emotional-Gain crimes, which is different than the literature's convention of grouping them into property crime and violent crime. The FBI's definition of property crime includes Burglary, Larceny-theft, Motor Vehicle Theft, and Arson (Federal Bureau of Investigation, 2018*a*). However, though not defined as a property crime, Robbery is also financially motivated, while Arson may not be. Therefore, the FBI's categorization of property crime is not optimal for examining how Poverty and Earnings affect crime. We modify the FBI's categorization by grouping together Robbery, Burglary, Larceny-theft, and Motor Vehicle Theft, and calling the group Financial-Gain crime. Similarly, the FBI's definition of violent crime includes Murder and non-negligent manslaughter, Rape, Robbery, and Aggravated Assault (Federal Bureau of Investigation, 2018*b*). However, Robbery is financially motivated, so it is excluded from our category of Emotional-Gain crime. Therefore, the category of Emotional-Gain crime includes Murder and non-negligent manslaughter, Rape, and Aggravated Assault.

Aggregating crimes at the CZ level has multiple advantages. When residents of one county make regular commutes across county lines, the crimes that they commit or experience are likely to cross county lines as well. If we examine the relationship between Poverty, Earnings and crime at the county level, we may encounter spurious effects of spillovers from one county to another. County borders can be complex, which makes residents of some counties live closer to the epicenters of other counties than to their own. This further increases the possibility of the crime spillover effect. Figure 2 provides an illustrative example. The city of Boston is in Suffolk County, but neighboring Cambridge and Somerville are in Middlesex County. It is not uncommon for residents of Boston to travel to Cambridge and Somerville on a regular basis, and vice versa. Moreover, Brookline is adjacent to Middlesex and Suffolk Counties, but is actually in Norfolk County. Criminals and victims both are likely to be in different counties within their Commuting Zone when crimes occur. Because CZs are defined based on actual commuting flow density, they are more likely to capture a complete picture of crime in the area.

Literature has established that not all serious crimes are reported to the UCR (Skogan, 1975, 1977, Ehrlich, 1977, Myers, 1980, Kennedy, 1988, Coleman and Moynihan, 1996, Levitt, 1998, Baumer and Lauritsen, 2010, Chalfin, 2015). For Financial-Gain crimes, some victims are not interested in reporting the crimes to the police as soon as they recover their items. For Emotional-Gain crimes, some victims may not be inclined to report the crimes, or may not be able to. Our results indicate that Poverty drives higher rates across both types of crime. If crime is under-reported in poorer Commuting Zones, then the actual effects of Poverty would be higher than they appear. This would mean that our estimates constitute a lower bound.

4 Standardized Crime Rates

In order to compare rates of crime of different kinds - such as Robbery and Murder - it is necessary to scale these rates relative to one another in a meaningful way. Intuitively, a single Murder is a larger incident than a single Robbery. But how much larger? It may be appropriate to cast the size of each criminal incident in terms of its impact - either in utility to the offender, or in disutility to the victim and society. For example, we might (very roughly) imagine that each Robbery yields a total utility cost to society of $300,000 on average, while each Murder yields an analogous cost of $10 million. However, such notions of impact are subjective, and can only be measured indirectly if at all.

To scale across different types of crime, we rely on the simple fact that more severe crimes tend to occur less frequently than milder crimes. The right-most column of Table 2 presents the grand total average rate for each type of crime in our data. The rarest (and intuitively most severe) type of crime is Murder, at 6.8 incidents per 100,000 people per year. Next is Rape, Robbery, Aggravated Assault, Motor Vehicle Theft, Burglary, and finally Larceny - at

2,383 incidents per 100,000 people per year. Following these grand average frequencies, we assume that, for example, a single Murder should be counted similarly as 2,328/6.8 = 350 Larcenies.

Formally, we designate a benchmark type of crime g = G to serve a numeraire for scaling crime rates. We believe Murder is the best choice for this benchmark, in part because typical Murder rates (per 100,000 people per year) fall in the range of 1-10, which is best for scaling purposes. The crime rate $\mathrm{C}_{\mathrm{g,t,z}}$, for any given type of crime g in time period t and location z, is scaled as:

$$\mathrm{C}_{\mathrm{g,t,z}} = (raw\ \mathrm{C}_{\mathrm{g,t,z}}) \cdot \mathrm{C}_{G}/(raw\ \mathrm{C}_{\mathrm{g}})\ ,\ \mathrm{C}_{G} = (raw\ \mathrm{C}_{G}) \tag{1}$$

Each *raw* rate is the actual number of incidents of crime type g per 100,000 people per year. The ratio of grand averages $\mathrm{C}_{G}/(raw\,\mathrm{C}_{\mathrm{g}})$ serves as a multiplier to make each rate comparable to a Murder rate. The central column of Table 2 provides these multipliers. For example, a single Murder is counted as 1 Standardized Incident, while a single Motor Vehicle Theft is counted as only 1/57 of a Standardized Incident.

5 Model

Like Becker (1968), we consider the utility of criminal offenders, who weigh their expected gains from crime against their expected losses from associated punishment - and the utility of the government (or Society), which weighs social losses from punishment against social losses from crime. Offenders set the crime rate, and the government sets a punishment policy. Our paper differs in that, where Becker stays in a more abstract space, we use concrete functional forms, which enable us to derive equilibrium solutions and estimate the associated parameters. We also focus much more on the role of expected legitimate (non-crime) earnings as a mechanism of interest, whereas Becker focuses more on risk preferences with respect to expected punishment.

It is worth taking note of the four motives at play in the model. The offender gains from crime and loses from punishment, while the government loses from both crime and from punishment.

	Crime	Punishment
Offender	$\uparrow$	$\downarrow$
Government	$\downarrow$	$\downarrow$

The first three of these motives are highly intuitive. It is the fourth (lower right) that may merit the most contemplation. The government (or Society) suffers from punishment for two reasons. First is simply financial cost - police, prosecution and incarceration cost money. Second, more deeply, is that the government internalizes (in principle) total social welfare, which includes the welfare of the prosecuted, both guilty and innocent. Note that this is entirely sufficient to yield, as a corollary, the intuitive idea that the innocent should not be prosecuted (or convicted). Because prosecution cannot deter crime except in so far as it is in response to crime, prosecution of the innocent has no useful function, and is pure loss. Prosecution of the guilty, though it may even be equally painful as prosecution of the innocent, can be justified only because it has a useful function as deterrent.

We begin by interpreting Becker's offender's utility in a more concrete functional form. An offender's expected net gain from criminal activity can be written as:

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - KC_i \ , \ \alpha \in [0,1] \tag{2}$$

where i is the offender (individual person), C_i is i's chosen frequency of crime, B is the expected benefit (gain to the offender) per crime, and K is the expected punishment per crime. We assume that the offender is risk-neutral with respect to punishment,[12] and use the power $(1-\alpha) \in [0,1]$ to represent Becker's supposition that offenders face decreasing marginal gain from crime.

We make two innovations upon equation (2). First, we assume that offenders, in addition to facing decreasing marginal gain from crime, face increasing marginal punishment:

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{1+\tilde{\alpha}} C_i^{1+\tilde{\alpha}}$$

For convenience, we suppose particularly that $\tilde{\alpha} = (1-\alpha)$:[13]

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{2-\alpha} C_i^{2-\alpha} \ , \ \alpha \in [0,1]$$

The rationale for increasing marginal punishment is that a more frequent offender is more likely to catch the focus of law enforcement and prosecution. Second, we assume that the offender's loss from criminal punishment is proportional to his expected legitimate (non-crime) earnings Y_i (plus a baseline $\underline{Y}$):[14]

$$U(C_i) = \frac{B}{1-\alpha} C_i^{1-\alpha} - \frac{K}{2-\alpha} C_i^{2-\alpha} (Y_i + \underline{Y}) \ , \ \alpha \in [0,1] \tag{3}$$

The rationale for this critical assumption is that people with higher expected earnings stand to lose more from criminal conviction.[15] In the case of incarceration, the offender would lose any earnings he would otherwise have gained if not behind bars. Even without incarceration, criminal records would still carry severe reputational damage, likely to impair the offender's future career. Note however that we leave the benefit B (as opposed to the cost K) portion

---

[12] Becker (1968) focuses largely on the distinction between the probability of conviction, and the severity of punishment conditional on condition; but this is not important for purposes. Our assumption of risk-neutrality renders the distinction inconsequential: K is the overall expected punishment per crime, a composite of the expected probability and the expected conditional punishment.

[13] In the limiting case of $\alpha = 1$, this is $U(C_i) = B \cdot log(C_i) - K \cdot C_i$.

[14] We assume $\underline{Y}$ is equal to 10% of average pre-tax earnings - see Appendix section A.3.

[15] Wealthier offenders' ability to pay for better lawyers may dampen this effect, but cannot negate or reverse it.

of the utility function independent of earnings. This independence can be questioned on the grounds that offenders may have decreasing marginal utility in total (legitimate + criminal) earnings, and therefore that criminal earnings may mean more in utility terms to lower legitimate-earning offenders. However, higher earning offenders may also have access to higher quality (such as "white-collar") opportunities for crime. Implicitly, we assume that these two factors offset one another.

The second agent in the model is the government, which chooses the level of punishment per crime K. (K is a composite of the probability of conviction, and the expected punishment conditional on conviction.) We assume that the government faces a loss function in which it suffers both from higher levels of K, and from higher crime rates C:[16]

$$L(K) = \vartheta^{-1} K^{\vartheta} + C(K) \ , \ \vartheta > 0 \tag{4}$$

Higher punishment per crime K is a loss to society for multiple reasons. Punishment is detrimental to offenders, whose welfare is part of the social welfare. Higher punishment per crime also implies costly government activities - more police, more prosecution, more incarceration - all of which fall heavily on the taxpayer. As the model is concerned, it is important only that higher K is undesirable.

Our analysis concerns different types of crime g, over different Commuting Zones z and time periods t. Written fully, equations (3) and (4) are:

Offender's Gain:

$$U(C_{i,g,t,z}) = \frac{B_{g,t,z}}{1-\alpha_g} C_{i,g,t,z}^{1-\alpha_g} - \frac{K_{g,t,z}}{2-\alpha_{g,}} C_{i,g,t,z}^{2-\alpha_g} (Y_{i,t,z} + \underline{Y}_t) \ , \ \alpha_g \in [0,1] \tag{5}$$

Government's Loss:

$$L(K_{g,t,z}) = \vartheta_g^{-1} K_{g,t,z}^{\vartheta_g} + C(K_{g,t,z}) \ , \ \vartheta_g > 0 \tag{6}$$

[16]Crime rates C are expressed in Standardized Incidents per 100,000 people per year - see Table 2.

The optimality conditions for each (given $\mathrm{U}''(\mathrm{C}^*_{\mathrm{i,g,t,z}}) < 0\ ,\ \mathrm{L}''_{\mathrm{t}}(\mathrm{K}^*_{\mathrm{g,t,z}}) > 0$)[17] are:

$$\mathrm{U}'(\mathrm{C}_{\mathrm{i,g,t,z}}) = 0\ ,\ \mathrm{L}'(\mathrm{K}_{\mathrm{g,t,z}}) = 0$$

That is, the offender i chooses a crime frequency $\mathrm{C}_{\mathrm{i,g,t,z}}$ to maximize U, and the government chooses an expected level of punishment per crime $\mathrm{K}_{\mathrm{g,t,z}}$ to minimize L.

For the offender, $\mathrm{U}'(\mathrm{C}_{\mathrm{i,g,t,z}}) = 0$ is equivalent to:

$$\mathrm{C}^*_{\mathrm{i,g,t,z}} = \mathrm{M}_{\mathrm{i,t,z}} \frac{\mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}}\ ,\ \mathrm{M}_{\mathrm{i,t,z}} = (\mathrm{Y}_{\mathrm{i,t,z}} + \underline{\mathrm{Y}}_{\mathrm{t}})^{-1} \tag{7}$$

The crime rate (in each Commuting Zone z in each time period t) then follows as the average crime frequency over individuals i:

Crime Rate:

$$\mathrm{C}_{\mathrm{g,t,z}} = \mathrm{N}^{-1}_{\mathrm{t,z}} \sum_{\mathrm{i} \in \mathrm{t,z}} \mathrm{C}^*_{\mathrm{i,g,t,z}} = \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}}\ ,\ \mathrm{M}_{\mathrm{t,z}} = \mathrm{N}^{-1}_{\mathrm{t,z}} \sum_{\mathrm{i} \in \mathrm{t,z}} (\mathrm{Y}_{\mathrm{i,t,z}} + \underline{\mathrm{Y}}_{\mathrm{t}})^{-1} \tag{8}$$

Plugging this in to (6), the government's loss function becomes:

$$\mathrm{L}(\mathrm{K}_{\mathrm{g,t,z}}) = \vartheta_{\mathrm{g}}^{-1} \mathrm{K}^{\vartheta_{\mathrm{g}}}_{\mathrm{g,t,z}} + \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{\mathrm{K}_{\mathrm{g,t,z}}} \tag{9}$$

The government's optimality condition ($\mathrm{L}'(\mathrm{K}_{\mathrm{g,t,z}}) = 0$) therefore is equivalent to:

$$\mathrm{K}^*_{\mathrm{g,t,z}} = (\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{1/(1+\vartheta_{\mathrm{g}})} \tag{10}$$

Plugging this back in to (8), the equilibrium crime rate is:

$$\mathrm{C}_{\mathrm{g,t,z}} = \frac{\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}}}{(\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{1/(1+\vartheta_{\mathrm{g}})}} = (\mathrm{M}_{\mathrm{t,z}} \mathrm{B}_{\mathrm{g,t,z}})^{\vartheta_{\mathrm{g}}/(1+\vartheta_{\mathrm{g}})} \tag{11}$$

[17] See Appendix section A.4 for derivation.

This result is intuitive. The equilibrium value of punishment per crime $K_{g,t,z}$ increases in each Commuting Zone z in response to the crime rate in each,[18] but without over-compensating. Therefore, punishment acts simply to dampen the crime rate. If punishment were constant over Commuting Zones, then the crime rate would be $C_{g,t,z} = M_{t,z}B_{g,t,z}$. Instead, due to punishment's response to the crime rate, the latter is dampened by the power $\vartheta_g/(1 + \vartheta_g)$, which is necessarily in the interval (0, 1) for $\vartheta_g > 0$.

We assume that the expected gain to offenders per crime $B_{g,t,z}$ is a Cobb-Douglas function of average earnings $Y_{t,z}$, and controls $X_{t,z}$ that include population density and racial demographics:

$$B_{g,t,z} = Y_{t,z}^{\beta_g^y} \cdot \{X_{t,z}^{\beta_g^x}\} \ , \ Y_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t) \tag{12}$$

The expected gain per crime $B_{t,z}$ is increasing in average earnings for the reason that higher earning people, as potential victims, possess more that is worth stealing. It is increasing in population density for the reason that higher interpersonal contact rates will present offenders with a more plentiful flow of potential crime opportunities, some of which will be better than others.

As additional controls $\{X_{t,z}\}$ we take, in the Baseline specification, the fractions of the local population who are Black and White, as well as regional fixed effects, and year by region effects. These Baseline controls are those which we are most confident to be exogenous: racial fractions are deeply inertial and persistent, because race is fully genetic, and embedded in particular locations by family ties. In a Fully Controlled specification, we include an additional five controls that we are less confident to be exogenous: the fractions of the local population who are Young (under 40), Jailed (currently incarcerated), College educated, Married, and Male. All of these demographic fractions should be interpreted (under the model) as forces that exert cultural influences on the perceived value of criminal activity, to all potential offenders in each location.

Plugging (12) into (11) and taking logs, we arrive at a closed-form linear expression for

[18] $K_{g,t,z}$ may be raised via heightened policy activity, or more aggressive prosecution.

the equilibrium log crime rate with respect to observed variables:

Equilibrium:

$$c_{g,t,z} = \Theta_g m_{t,z} + \Theta_g \beta_g^y y_{t,z} + \{\Theta_g \beta_g^x x_{t,z}\} \tag{13}$$

$$c_{g,t,z} = log(C_{g,t,z}) \ , \ m_{t,z} = log(M_{t,z}) \ , \ y_{t,z} = log(Y_{t,z}) \ , \ x_{t,z} = log(X_{t,z})$$

$$\Theta_g = \vartheta_g/(1+\vartheta_g)$$

where $Y_{t,z}$ is the average earnings (in Commuting Zone z in time period t), $M_{t,z}$ is the average inverse earnings, and $X_{t,z}$ are controls, including population density and racial demographics. We estimate this equation using Two-Stage Least Squares, instrumenting for $m_{t,z}$ and $y_{t,z}$ with Bartik instruments. The coefficient on $m_{t,z}$ recovers $\Theta_g = \vartheta_g/(1+\vartheta_g)$. The coefficient on $y_{t,z}$ recovers $\Theta_g\beta_g^y$, which implies the value of β_g^y given that for Θ_g; and so on for each β_g^x.

$Y_{t,z}$ and $M_{t,z}$ are the explanatory variables of interest. $Y_{t,z}$ represents the upper tail of the earnings distribution, while $M_{t,z}$ represents poverty (the lower tail). Although $M_{t,z}$ is the average inverse of earnings, it is *not* the inverse of $Y_{t,z}$.

$$Y_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t)$$

$$M_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (Y_{i,t,z} + \underline{Y}_t)^{-1} > Y_{t,z}^{-1}$$

The latter inequality holds unless all values of $Y_{i,t,z}$ are identical, in which case $M_{t,z} = Y_{t,z}^{-1}$. Given any variation in individual earnings ($Y_{i,t,z}$), then $M_{t,z} > Y_{t,z}^{-1}$, because $M_{t,z}$ places disproportionate weight on the lower tail of the distribution. As such $M_{t,z}$, which arises organically as a central variable in the model from the assumption that lower earning people have less to lose in the context of criminal punishment, is ideal as a measure of poverty. $Y_{t,z}$ is merely the average, but places disproportionate weight on the upper tail in the sense that it is affected by the right-skew characteristic of earnings distributions.[19] Thus, $\mathrm{Y}_{\mathrm{t,z}}$ and $\mathrm{M}_{\mathrm{t,z}}$ capture the upper and lower tails in juxtaposition.

6 Identification

Two main identification hurdles arise with regard to estimating (13): reverse causality, and omitted explanatory variables. Our responses to each in turn are: Shift-Share (Bartik) instruments, and controls that include regional fixed effects and year by region effects.

6.1 Shift-Share (Bartik) Instruments

In equation (13), the response (left hand side) variables are crime rates, while the main explanatory (right hand side) variables of interest are average earnings, and average inverse earnings. The basic economic idea coinciding with equation (13) is that the distribution of legitimate (non-crime) earnings plays a role (or multiple roles) in motivating people to engage in crime. However, crime rates may also have reverse effects on the distribution of legitimate earnings. Unfavorable crime rates may motivate individuals with the means to move to do so, thereby dampening what would otherwise have the upper tail of the earnings distribution in any given Commuting Zone. Amongst the poor, crime may contribute to negative feedback loops, pushing more people into crime.

To get around any such effects of crime on earnings, we construct a set of Shift-Share (Bartik) instruments.[20] The instrument, for each variable, interacts the lagged fractional breakdown of employment opportunities in each location z across occupational categories, with the corresponding vector of national averages for each category. This arrives at a measure for each earnings variable that is decoupled from the effects of contemporary local crime rates in each location z, depending rather only on the historical structure of the economy in each.

[19] $\mathrm{Y}_{\mathrm{t,z}}$ places disproportionate weight on the upper tail relative to, for example, the median of $\mathrm{Y}_{\mathrm{i,t,z}}$, or the average of $log(\mathrm{Y}_{\mathrm{i,t,z}})$.

[20] The name Bartik comes from Bartik (1991), but refers to a broad class of instruments. See Bound and Holzer (2000) and Diamond (2016) for examples.

The component of the instrument that is specific to each location z is what we call the Occupational Profile (OP). The OP of each Commuting Zone z, in each time period t, is a vector of 14 fractions summing to one, giving the occupational breakdown of employment across each in z. That is, the OP for t,z is the vector,

$$\left\{ \frac{\mathrm{NE}_{\mathrm{o,t,z}}}{\mathrm{NE}_{\mathrm{t,z}}} \right\}_{\mathrm{o}}$$

where $\mathrm{NE}_{\mathrm{o,t,z}}$ is the count of people employed in Occupation o, and $\mathrm{NE}_{\mathrm{t,z}}$ is the count of people employed in any Occupation. To arrive at the instrument for Earnings, the OP is interacted with the vector,

$$\left\{ \mathrm{Y}_{\mathrm{o,t}} \right\}_{\mathrm{o}}$$

where $\mathrm{Y}_{\mathrm{o,t}}$ is the national average post-tax earnings of people who identify Occupation o as their habitual occupation (even if not employed).

We assume that OPs are a first mover of the economic system under study, arising from a combination of exogenous geographical, historical, and technological factors. In other words, OPs do not depend on crime rates. For example: although unfavorable crime rates may influence average earnings in Chicago via out-migration or societal trauma, crime does not influence the probability that a worker in Chicago, if employed, will be employed in finance. In case this is considered questionable, we lag OPs by one time period. The one period lag allows OPs to update for relevance to the nature of economic activity in the current time period, while remaining clearly causally prior to crime in the current period. The instrument for Earnings therefore is,

$$\tilde{\mathrm{y}}_{\mathrm{t,z}} = log(\tilde{\mathrm{Y}}_{\mathrm{t,z}}) \ , \ \tilde{\mathrm{Y}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \frac{\mathrm{NE}_{\mathrm{o,t-1,z}}}{\mathrm{NE}_{\mathrm{t-1,z}}} \cdot \mathrm{Y}_{\mathrm{o,t}} \tag{14}$$

where t = (0, 1, 2, 3, 4) indicate survey years (1980, 1990, 2000, 2010, 2016) respectively.

The instrument for Poverty is analogous,

$$\tilde{m}_{t,z} = log(\tilde{M}_{t,z}) \ , \ \tilde{M}_{t,z} = \sum_{o} \frac{NE_{o,t-1,z}}{NE_{t-1,z}} \cdot M_{o,t} \tag{15}$$

The instrument given by (15) is novel, but follows by the same logic as (14), which is a more standard Shift-Share instrument. $M_{o,t}$, like $Y_{o,t}$, is a summary of the distribution of Occupation o specific (but not location z specific) earnings at the national level, so is exogenous for the same reasons.

6.2 Fixed Effects

As is almost always the case in any real setting, the potential for omitted unobservables to yield bias is the truest and most intractable identification issue. We rely on a standard method that is meant to (plausibly/perhaps) absorb all relevant unobserved information, namely, fixed effects. Because not all potentially relevant factors are of a nature that is fixed over time, we use region-by-year effects as well regional fixed effects. For example, the weather in New York is about equally worse than the weather in Florida regardless of the time period, but the same cannot always be said of things like cultural influences. We use the nine regional divisions used by the Census,[21] and further split each of these into their urban and rural Commuting Zones,[22] for a total of 18 regions. We include an effect for each of these, in each of the 5 decadal time periods, for a total of 90 effects.

7 Descriptive Statistics

This paper uses a host of socioeconomic and demographic factors as explanatory variables in determining equilibrium crime rates. Table 4 summarizes these explanatory and control variables, scaled for readability. Means are in rows without parentheses, while the standard deviation for each is in parentheses in the row beneath. The percentage of Whites in the

[21] These nine regional divisions have the benefit of being about the same size as one another.
[22] See Appendix section A.1.

mean Commuting Zone has decreased over the course of 1980-2016, as has the percentage of Young (under 40) people, and the Marriage rate. Meanwhile, the percentage of Blacks, the incarceration (Jailed) rate, and the percentage of people who are College educated in the mean CZ have all increased. All of the above reflect, but are not equivalent to, corresponding changes in the national mean of each variable.[23]

It should be noted that although mean values of all the variables as discussed in the previous paragraph have changed over time, this does not make a difference in the model estimation. In the same vein, adjusting earnings figures for inflation would not make any difference. This is because, in logarithmic form, all scaling factors become additive intercepts, which in the model estimation are absorbed in time period effects fixed effects.

Unlike the scaling of variables above, the scaling of crime rates is extremely important. Table 3 summarizes the distribution of (CZ average) crime rates, expressed in Standardized Incidents as defined in Section 4 and Table 2. Financial-Gain crime is Robbery, Burglary, Larceny, and Motor Vehicle Theft. Emotional-Gain crime is Murder, Rape, and Aggravated Assault. Both types of crime peaked in 1990. More concretely still, Table 5 lists individual Commuting Zones by their crime rates in 2010. One illustrative example is Detroit - 7th in Emotional-Gain crimes, but not in the top 15 for Financial-Gain crimes. Detroit has a pronounced lower tail of earnings (Poverty) - so has high crime rates in general. But it does not have a pronounced upper tail of earnings. Therefore in Financial-Gain crimes - which are driven by both Poverty and high Earnings - Detroit is overtaken by some higher earning CZs such as Houston and San Francisco.

8 Results

Our main parameters of interest are ϑ_{g} and β^{y}_{g}, for each of Emotional-Gain crimes ($\mathrm{g} = E$) and Financial-Gain crimes ($\mathrm{g} = F$). ϑ identifies the role of Poverty in driving crime,[24]

[23]The national mean would weight each individual equally. The values in Table 4 are means of CZ means, weighting each CZ equally, which in effect gives more weight to individuals in lower populated CZs.

[24]ϑ also pins down the government's loss function, by weighing the harm arising from higher punishment per crime against the harm arising from crime directly.

while β^y identifies the role of high Earnings. Our central finding, both theoretically and empirically, is that β^y is higher for Financial-Gain crimes than it is for Emotional-Gain crimes ($\beta^y_F > \beta^y_E$), while ϑ is about the same for both types of crime ($\vartheta_F = \vartheta_E$). Table 8 encapsulates these main results. Additionally, we find that not accounting for reverse causality would underestimate both parameters, for both types of crime, often by more than 50%.[25]

The distinct roles of Poverty and higher Earnings in driving crime are such that the former motivates all types of crime, whereas the latter motivates only (or principally) crimes coinciding with financial gain to the offender. The role of Poverty enters into the model via the assumption that the offender's disutility in the event of criminal punishment is proportional to his legitimate (non-crime) earnings.[26] This assumption, combined with optimality and equilibrium conditions, yields the average inverse of earnings (the variable $m_{t,z}$) as a structural representative of Poverty. Therefore the parameter, ϑ, that is associated with $m_{t,z}$ should attain similar estimates across both categories of crime (Financial-Gain and Emotional-Gain).

The upper tail of the earnings distribution, unlike Poverty (the lower tail), should not motivate both types of crime commensurately. High Earnings enter the model via the assumption that higher earning people possess more that is worth stealing, and therefore present more fruitful criminal opportunities to potential offenders. This logic applies far more to Financial-Gain crimes (Robbery, Burglary, Larceny, Motor Vehicle Theft) than it does to Emotional-Gain crimes (Murder, Rape, Aggravated Assault). Therefore the parameter β^y, which captures the role of high Earnings, should be higher for the Financial-Gain crimes than it is for Emotional-Gain crimes.

The parameter estimates summarized in Table 8 corroborate that the role of high Earnings (β^y) is higher for Financial-Gain crimes (1.185) than for Emotional-Gain crimes (0.525), while the role of Poverty, $\Theta = \vartheta/(1 + \vartheta)$, is about the same (0.788 versus 0.750) across

[25]This is apparent in Table 7.

[26]That is, people with lower legitimate earnings have less to lose the context of criminal punishment, and therefore more reason (on balance) to engage in crime.

both categories of crime. These estimates are taken from Table 7's lower panel, the Fully Controlled specification. The Baseline specification (Table 7's upper panel) suggests that β^y for Emotional-Gain crimes is not statistically significant at all, which would corroborate our theory even more strongly.

One reason that β^y may be positive for Emotional-Gain crimes (albeit not as large as that for Financial-Gain crimes) is that the two types of crime are linked via gang activity. Murder, for example, although not intrinsically associated with financial gain for the offender, may be deployed strategically by criminal enterprises that are driven ultimately by illegal revenue streams. A second reason is that the presence of high Earnings may inspire envy amongst potential offenders, thus enhancing the emotional impetus for crime.

9 Counterfactual Toughness on Crime

The most important parameter in our model, ϑ, is also one for which we receive some of the most consistent estimates across different specifications. By definition, ϑ is the government's distaste for criminal punishment: a lower value of ϑ indicates a tougher stance on crime. By derivation, ϑ also coincides with the elasticity of crime with respect to Poverty; and is a factor in the elasticity of crime with respect to all other motivating factors as well.

Because ϑ is a parameter of the government's loss function, it is in essence a policy choice. We simulate counterfactual changes in the value of ϑ in order to examine the tradeoff it faces between crime and punishment. It follows[27] from the equilibrium crime rate (11) and equilibrium level of expected punishment per crime (10) that,

$$\check{\check{\mathrm{C}}}_{\mathrm{z}} = \mathrm{C}_{\mathrm{z}}^{\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

$$\check{\check{\mathrm{K}}}_{\mathrm{z}} = \mathrm{C}_{\mathrm{z}}^{-\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

$$\check{\check{\vartheta}} := (\check{\vartheta} - \vartheta)/\vartheta \ , \ \check{\check{\mathrm{C}}}_{\mathrm{z}} := (\check{\mathrm{C}}_{\mathrm{z}} - \mathrm{C}_{\mathrm{z}})/\mathrm{C}_{\mathrm{z}} \ , \ \check{\check{\mathrm{K}}}_{\mathrm{z}} := (\check{\mathrm{K}}_{\mathrm{z}} - \mathrm{K}_{\mathrm{z}})/\mathrm{K}_{\mathrm{z}}$$

[27]See Appendix section A.5 for derivation.

where $\check{\vartheta}$ is a counterfactual theta value of policy choice, and $\check{\mathrm{C}}_{\mathrm{z}}$ and $\check{\mathrm{K}}_{\mathrm{z}}$ are the resulting counterfactual equilibrium crime rate and level of expected punishment per crime, respectively. For example, a value of $\check{\check{\vartheta}} = 0.5$ would mean that the counterfactual policy choice is to increase ϑ by 50%. $\check{\check{\mathrm{K}}}_{\mathrm{z}} = -0.3$ would mean that the resulting equilibrium expected punishment per crime would be 30% lower than in the factual equilibrium, and $\check{\check{\mathrm{C}}}_{\mathrm{z}} = 0.2$ that the resulting crime rate would be 20% higher.

As ϑ represents the government's distaste for punishment, a higher value of ϑ must coincide with a loosening of punishment, and consequently higher equilibrium crime rates. The two right columns of Table 9 summarize the consequences of a 20% increase in ϑ_{g} for each type of crime g. For Financial-Gain crimes ($\mathrm{g} = F$), this would result in a 6.2% decrease in expected punishment per crime in the median Commuting Zone, and a corresponding 6.6% increase in crime. A 20% decrease in ϑ, on the other hand, would result in 9.2% increase in expected punishment per crime, and a 8.4% decrease in crime.

Being an abstract policy disposition, it is not obvious what a 20% change in the value of ϑ would look like concretely. However, it links together two concrete outcomes - the crime rate (C), and expected level of punishment per crime (K). A 9.2% increase in K can be interpreted either as a 9.2% increase in the probability of conviction, or a 9.2% increase in the severity of punishment conditional on conviction, with the other held constant. Our results indicate that, for example, in order to achieve an 8.4% decrease in crime, the median Commuting Zone would need to increase the severity of its criminal sentences (or its conviction probability) by 9.2%. In other words, there are no particularly great options in the tradeoff between crime and punishment.

10 Discussion

Our results provide concrete evidence for the benefits of reducing inequality, and especially poverty. Many current poverty alleviation efforts exist in the US. For example, summer youth employment programs (SYEP) are known to alleviate both poverty and crime (Modestino, 2019, Davis and Heller, 2020, Kessler et al., 2021). Researchers have found that SYEP increase employment outcomes in a subset of youths. They have also found suggestive evidence that the programs improved youths' conflict resolution skills, including self-regulation and ability to respond positively to criticism. Additional improvements may include peer networks and income. Currently these programs are conducted in Boston, Chicago, and New York: more widespread adoption would certainly be beneficial. Another method is Moving to Opportunity (MTO) (Chetty, Hendren and Katz, 2016). Researchers found that MTO helps children before age 13 in areas including college attendance, earnings, and single parenthood rates.

Another important line of work highlights the vicious cycle between poverty and productivity decline (Banerjee and Mullainathan, 2008, Kaur et al., 2021, Duquennois, 2022). Increasing the savings rate of low-income individuals can mitigate this productivity decline. There are two obstacles to this solution. First, low-income individuals tend to be more present-biased before payday (Carvalho, Meier and Wang, 2016). Second, low-income households have low participation rate in 401(k) (Poterba, Venti and Wise, 2000). To overcome these problems, companies should automatically enroll their low-income employees in 401(k) (Thaler and Benartzi, 2004). Bhargava and Manoli (2015) have found mailing, simplification, and heightening salience of benefits increase take-up of EITC benefits. Hoynes, Schanzenbach and Almond (2016) have found that access to food stamps in childhood increases health outcomes and, for women, later-life economic self-sufficiency.

Finally, taxes can play a large role in social welfare. A negative tax on the poor can help overcome hyperbolic discounting and myopia (Farhi and Gabaix, 2020). Other researchers have suggested overall tax reforms to reduce inequality (Ales, Kurnaz and Sleet, 2015, Altig and Carlstrom, 1999). The above methods of reducing inequality and poverty, amongst others, can increase legitimate earnings of the poor and in turn reduce crime.

11 Conclusion

Much work in economics has argued that earnings inequality plays a role in driving crime. This paper is novel in distinguishing the role of the lower tail of the earnings distribution (Poverty) in driving crime, from that of the upper tail. We find that the roles played by these two tails driving crime are different in kind. Poverty results in individual offenders who have less to lose from criminal punishment, and are therefore less averse to engaging in all forms of crime. High Earnings heighten the expected benefit to offenders per crime, but only for crimes that yield financial gain to the offender. As a result, high Earnings drive only a subset of crimes, while Poverty results in higher rates of all types of crime.

We develop a new model that articulates how Poverty and Earnings become factors in determining equilibrium crime rates. The model has two players - the criminal offender, and the government. The offender maximizes his utility by choosing his frequency of crime. The government minimizes social loss by choosing the level of expected punishment per crime. In equilibrium, the crime rate in each Commuting Zone is a function of Poverty, Earnings, and demographic factors.

In order to estimate the model, we construct a comprehensive panel of Census, American Community Survey, and FBI Uniform Crime Reporting data, covering the entire United States (722 Commuting Zones) from 1980-2016. To correct for reverse causality, we deploy novel Shift-Share instruments for different parts of the earnings distribution. The instruments assume that crime impacts within-industry earnings in each Commuting Zone, but not the historical industrial makeup of each CZ, nor within-industry earnings distributions at the national level.

We find that higher Earnings significantly increase rates of Financial-Gain crimes (Robbery, Burglary, Larceny, Motor Vehicle Theft), while Poverty significantly increases rates both of Financial-Gain and of Emotional-Gain crimes (Murder, Rape, Aggravated Assault). Removing reverse causality results in substantially higher parameter estimates, in some cases by more than double.

References

Ales, Laurence, Musab Kurnaz, and Christopher Sleet. 2015. "Technical change, wage inequality, and taxes." *American Economic Review*, 105(10): 3061–3101.

Altig, David, and Charles T Carlstrom. 1999. "Marginal tax rates and income inequality in a life-cycle model." *American Economic Review*, 89(5): 1197–1215.

Autor, David, and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David, David Dorn, and Gordon Hanson. 2013. "When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men." *American Economic Review*, 103(5): 1553–1597.

Banerjee, Abhijit V, and Sendhil Mullainathan. 2008. "Limited attention and income distribution." *American Economic Review*, 98(2): 489–93.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Basu, Rounaq, and Joseph Ferreira. 2021. "Sustainable mobility in auto-dominated Metro Boston: Challenges and opportunities post-COVID-19." *Transport Policy*, 103: 197–210.

Baumer, Eric P, and Janet L Lauritsen. 2010. "Reporting crime to the police, 1973–2005: a multivariate analysis of long-term trends in the National Crime Survey (NCS) and National Crime Victimization Survey (NCVS)." *Criminology*, 48(1): 131–185.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment." *American Economic Review*, 105(11): 3489–3529.

Bound, John, and Harry Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

Brush, Jesse. 2007. "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.

Burdett, Kenneth, Ricardo Lagos, and Randall Wright. 2003. "Crime, inequality, and unemployment." *American Economic Review*, 93(5): 1764–1777.

Carvalho, Leandro S, Stephan Meier, and Stephanie W Wang. 2016. "Poverty and economic decision-making: Evidence from changes in financial resources at payday." *American economic review*, 106(2): 260–84.

Chalfin, Aaron. 2015. "The long-run effect of Mexican immigration on crime in US cities: Evidence from variation in Mexican fertility rates." *American Economic Review*, 105(5): 220–25.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review*, 106(4): 855–902.

Choe, Jongmook. 2008. "Income inequality and crime in the United States." *Economics Letters*, 101(1): 31–33.

Coleman, Clive, and Jenny Moynihan. 1996. *Understanding crime data: Haunted by the dark figure.* Vol. 120, Open University Press Buckingham.

Davis, Jonathan MV, and Sara B Heller. 2020. "Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs." *Review of Economics and Statistics*, 102(4): 664–677.

Demombynes, Gabriel, and Berk Özler. 2005. "Crime and local inequality in South Africa." *Journal of development Economics*, 76(2): 265–292.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524.

Dorn, David. 2009. "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD diss. University of St. Gallen.

Duquennois, Claire. 2022. "Fictional money, real costs: Impacts of financial salience on disadvantaged students." *American Economic Review*, 112(3): 798–826.

Ehrlich, Isaac. 1973. "Participation in illegitimate activities: A theoretical and empirical investigation." *Journal of political Economy*, 81(3): 521–565.

Ehrlich, Isaac. 1975. "The Deterrent Effect of Capital Punishment: A Question of Life and Death." *American Economic Review*, 65(3): 397–417.

Ehrlich, Isaac. 1977. "Capital punishment and deterrence: Some further thoughts and additional evidence." *Journal of Political Economy*, 85(4): 741–788.

Ehrlich, Issac. 1970. "Participation in Illegitimate Activities: An Economic Analysis." PhD diss. Columbia University.

Enamorado, Ted, Luis F López-Calva, Carlos Rodríguez-Castelán, and Hernán Winkler. 2016. "Income inequality and violent crime: Evidence from Mexico's drug war." *Journal of Development Economics*, 120: 128–143.

Engelhardt, Bryan, Guillaume Rocheteau, and Peter Rupert. 2008. "Crime and the labor market: A search model with optimal contracts." *Journal of Public Economics*, 92(10-11): 1876–1891.

Fajnzylber, Pablo, Daniel Lederman, and Norman Loayza. 2002. "Inequality and violent crime." *The journal of Law and Economics*, 45(1): 1–39.

Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal taxation with behavioral agents." *American Economic Review*, 110(1): 298–336.

Federal Bureau of Investigation. 1980. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 1977-1983." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/8703` (accessed August 13, 2021).

Federal Bureau of Investigation. 1990. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 1990." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/9785` (accessed August 12, 2021).

Federal Bureau of Investigation. 2000. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2000." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/3451` (accessed August 12, 2021).

Federal Bureau of Investigation. 2010. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2010." United States Department of Labor. `https://www.icpsr.umich.edu/web/NACJD/studies/33526` (accessed August 12, 2021).

Federal Bureau of Investigation. 2016. "Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 2016." United States Department of Labor. https://www.icpsr.umich.edu/web/NACJD/studies/37059 (accessed August 13, 2021).

Federal Bureau of Investigation. 2018*a*. "*UCR Property Crime Definition.*" https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/topic-pages/property-crime (Accessed: September 28, 2021).

Federal Bureau of Investigation. 2018*b*. "*UCR Violent Crime Definition.*" https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/topic-pages/violent-crime (Accessed: September 28, 2021).

Florida, Richard, and Charlotta Mellander. 2016. "Rise of the startup city: The changing geography of the venture capital financed innovation." *California Management Review*, 59(1): 14–38.

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

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics*, 132(4): 1693–1754.

Hipp, John R. 2007. "Income inequality, race, and place: Does the distribution of race and class within neighborhoods affect crime rates?" *Criminology*, 45(3): 665–697.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." *American Economic Review*, 106(4): 903–34.

International Centre for Prison Studies. 2017-2018. "World Prison Brief Data." https://www.prisonstudies.org/world-prison-brief-data (accessed February 28, 2022).

Kang, Songman. 2016. "Inequality and crime revisited: Effects of local inequality and economic segregation on crime." *Journal of Population Economics*, 29(2): 593–626.

Kaur, Supreet, Sendhil Mullainathan, Suanna Oh, and Frank Schilbach. 2021. "Do Financial Concerns Make Workers Less Productive?" National Bureau of Economic Research.

Kelly, Morgan. 2000. "Inequality and crime." *Review of economics and Statistics*, 82(4): 530–539.

Kennedy, Leslie W. 1988. "Going it alone: Unreported crime and individual self-help." *Journal of Criminal Justice*, 16(5): 403–412.

Kessler, Judd B, Sarah Tahamont, Alexander M Gelber, and Adam Isen. 2021. "The Effects of Youth Employment on Crime: Evidence from New York City Lotteries." National Bureau of Economic Research.

Levitt, Steven D. 1998. "The relationship between crime reporting and police: Implications for the use of uniform crime reports." *Journal of quantitative criminology*, 14(1): 61–81.

Matarazzo, Thomas, Mohammad Vazifeh, Shamim Pakzad, Paolo Santi, and Carlo Ratti. 2017. "Smartphone data streams for bridge health monitoring." *Procedia engineering*, 199: 966–971.

Modestino, Alicia Sasser. 2019. "How do summer youth employment programs improve criminal justice outcomes, and for whom?" *Journal of Policy Analysis and Management*, 38(3): 600–628.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3): 173–196.

Myers, Samuel L. 1980. "Why are Crimes Underreported? What is the Crime Rate? Does it" Really" Matter?" *Social Science Quarterly*, 61(1): 23–43.

Poterba, James M, Steven F Venti, and David A Wise. 2000. "Saver behavior and 401 (k) retirement wealth." *American Economic Review*, 90(2): 297–302.

Quadrini, Vincenzo. 2020. "The impact of industrialized countries' monetary policy on emerging economies." *IMF Economic Review*, 68(3): 550–583.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN.

Skogan, Wesley G. 1975. "Measurement problems in official and survey crime rates." *Journal of criminal justice*, 3(1): 17–31.

Skogan, Wesley G. 1977. "Dimensions of the dark figure of unreported crime." *Crime & Delinquency*, 23(1): 41–50.

Stolzenberg, Lisa, David Eitle, and Stewart J D'alessio. 2006. "Race, economic inequality, and violent crime." *Journal of Criminal Justice*, 34(3): 303–316.

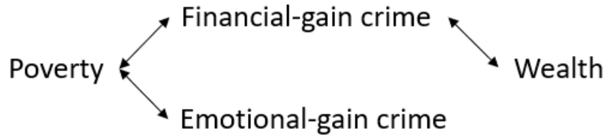
Thaler, Richard H, and Shlomo Benartzi. 2004. "Save more tomorrow™: Using behavioral economics to increase employee saving." *Journal of political Economy*, 112(S1): S164–S187.

Tolbert, Charles M, and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update."

Witte, Ann Dryden. 1980. "Estimating the economic model of crime with individual data." *The quarterly journal of economics*, 94(1): 57–84.

Tables and Figures

Figure 1: Variable Relationsips



Poverty motivates (+) crime, but crime can exacerbate (+) Poverty. OLS results are therefore biased downward. High Earnings motivate (+) Financial-Gain crime, and crime can diminish (−) Earnings, such as via out-migration. The resulting direction of bias in OLS can be either upward or downward for this latter case.

Table 1: Incarceration Rates in Industrialized Countries

Country	Incarceration Rate
United States	642
New Zealand	214
Australia	172
United Kingdom	139
Spain	128
Portugal	125
France	114
Canada	112
Italy	99
Austria	98
Greece	93
Belgium	90
Switzerland	82
Ireland	80
Germany	77
Denmark	65
Norway	65
Netherlands	63
Sweden	63
Finland	53
Japan	40

Rates are per 100,000 population in 2017-2018. Data are from Quadrini (2020) and the International Centre for Prison Studies (2017-2018).

Table 2: Aggregate Crime Rates and Conversions

Crime Type (g)	(*raw* C_g)/C_G	(*raw* C_g)
F		
Robbery	1/24.34	164.99
Burglary	1/132.4	897.91
Larceny	1/351.6	2383.8
Motor Vehicle Theft	1/57.05	386.78
E		
Murder (G)	1	6.7796
Rape	1/5.099	34.571
Aggravated Assault	1/47.77	323.85

The right column (*raw* C_g) provides the grand average rate of each type of crime (g) from 1980-2016 over the whole nation, expressed in incidents per 100,000 people per year. The central column provides the multipliers that convert any raw rates for each type of crime into expression in Standardized Incidents. In other words, a Standardized Incident is 1 Murder, 57 Motor Vehicle Thefts, or so on. These conversions are meant to adjust for the severity of each type of crime - resting on the assumption that more severe crimes occur proportionally less frequently in aggregate than milder crimes do. Crime type (g $= F$) is the average for Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), while (g $=$ E) is the average for Emotional-Gain crimes (Murder, Rape, and Aggravated Assault). Original crime rates must be converted into Standardized Incidents before averages F and E are computed.

Table 3: Descriptive Statistics: Crime

Variable	Statistic	1980	1990	2000	2010	2016
$C_{F,\mathrm{t,z}}$	*mean*	5.663	7.734	5.989	4.666	2.904
	stdev	3.467	3.747	2.436	1.586	1.460
	p10	2.400	3.820	2.957	2.541	1.418
	p50	4.641	6.786	5.792	4.621	2.569
	p90	10.29	12.23	8.925	6.642	4.980
$C_{E,\mathrm{t,z}}$	*mean*	5.554	7.693	5.995	5.155	5.849
	stdev	3.033	3.687	2.289	1.671	5.230
	p10	2.257	3.182	3.186	3.060	3.006
	p50	4.874	7.822	5.962	4.995	5.385
	p90	9.567	12.74	9.200	7.452	9.000

$C_{F,\mathrm{z}}$ is the rate of Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and $C_{E,\mathrm{z}}$ is the rate of Emotional-Gain crimes (Murder, Rape, and Aggravated Assault), in each Commuting Zone, z. Crime rates are expressed in Standardized Incidents per 100,000 people per year. A Standardized Incident is 1 Murder, 24 Robberies, 57 Motor Vehicle Thefts, or so on - see Table 2. t restricts by year. Included are only the 250 largest Commuting Zones by population.

Table 4: Descriptive Statistics: Earnings and Demographics

Variable	1980	1990	2000	2010	2016
$M_{t,z}$ ($\$^{-1}10^{-4}$)	3.803	2.057	1.251	1.053	0.780
	(0.784)	(0.351)	(0.181)	(0.129)	(0.086)
$Y_{t,z}$ ($\$10^3$)	8.551	14.51	21.16	27.96	33.63
	(1.075)	(1.881)	(2.743)	(3.075)	(3.599)
$Q_{t,z}$ (mi^{-2})	85.86	91.36	104.1	110.4	116.2
	(188.2)	(245.4)	(275.9)	(248.1)	(299.8)
$P_{t,z}^{Black}$ (%)	7.529	7.538	8.087	8.657	8.975
	(9.110)	(9.188)	(9.813)	(10.30)	(10.23)
$P_{t,z}^{White}$ (%)	87.81	87.29	83.71	83.45	82.20
	(14.67)	(9.411)	(10.31)	(10.47)	(10.47)
$P_{t,z}^{Young}$ (%)	58.89	58.70	51.76	47.42	48.83
	(8.615)	(3.000)	(3.339)	(3.642)	(3.276)
$P_{t,z}^{Jailed}$ (%)	0.758	1.170	1.958	2.109	2.141
	(0.586)	(0.990)	(1.758)	(1.990)	(1.736)
$P_{t,z}^{College}$ (%)	5.316	15.01	17.84	20.22	21.75
	(1.953)	(3.989)	(4.724)	(5.249)	(5.767)
$P_{t,z}^{Married}$ (%)	67.01	64.82	61.56	53.82	50.57
	(10.08)	(4.452)	(3.953)	(4.321)	(3.962)
$P_{t,z}^{Male}$ (%)	48.07	49.33	50.04	50.52	50.56
	(6.588)	(1.135)	(1.278)	(1.241)	(1.262)

Listed are (unweighted) means of CZ means, not national means. For example, the means for the fraction *White* are higher than the national fraction of Whites would be, because Whites are relatively concentrated in lower population CZs. Standard deviations (of the CZ means) are in parentheses. Y_z is average post-tax earnings, and M_z is average post-tax inverse earnings, per year amongst age 18-64 people in the labor force in each Commuting Zone, z. Q_z is population density. Demographic fractions P are restricted to age 18-64 people who are (except in the case of *Jailed*) not institutionalized. *Young* indicates age 18-40. *Jailed* means currently incarcerated. *College* means at least four years of higher education completed. t restricts by year.

Table 5: Highest and Lowest Commuting Zones in 2010 Crime Rates

	$C_{F,t,z}$	CZ	State	$C_{E,t,z}$	CZ	State
max	9.280	Memphis	TN	10.72	New Orleans	LA
2	9.119	Columbus	GA	10.52	Memphis	TN
3	8.462	Little Rock	AR	9.180	Little Rock	AR
4	8.268	Fayetteville	NC	8.900	Springfield	IL
5	8.093	Florence	SC	8.792	Redding	CA
6	8.028	Houston	TX	8.422	Albuquerque	NM
7	7.727	Jackson	MS	8.359	Detroit	MI
8	7.717	Augusta	GA	8.358	Columbia	SC
9	7.689	Columbus	OH	8.336	Baton Rouge	LA
10	7.684	Bakersfield	CA	8.245	Shreveport	LA
11	7.634	Indianapolis	IN	8.242	Tallahassee	FL
12	7.366	Yakima	WA	8.165	Corpus Christi	TX
13	7.281	San Antonio	TX	8.103	Baltimore	MD
14	7.201	San Francisco	CA	8.103	Florence	SC
15	7.137	Fresno	CA	8.049	Lubbock	TX
60	5.901	Orlando	FL	6.291	Macon	GA
61	5.890	Chicago	IL	6.162	Lansing	MI
p75	5.882	Toledo	OH	6.151	Ocala	FL
63	5.870	Jacksonville	FL	6.141	Fayetteville	NC
64	5.860	Cincinnati	OH	6.137	Beaumont	TX
123	4.655	Killeen	TX	5.038	Athens	GA
124	4.655	Hickory	NC	5.037	Pueblo	CO
p50	4.644	Colorado Springs	CO	5.008	Johnson City	TN
126	4.598	San Jose	CA	4.981	Columbus	GA
127	4.577	Jolpin	MO	4.971	Phoenix	AZ
185	3.445	Columbia	MO	4.075	Boise	ID
186	3.407	Cedar Rapids	IA	4.055	Gastonia	NC
p25	3.399	Albany	NY	4.037	Portland	OR
188	3.379	Lima	OH	4.025	St. Cloud	MN
189	3.377	Kennewick	WA	4.012	San Jose	CA
246	2.016	Eau Claire	WI	2.141	Findlay	OH
247	1.888	State College	PA	1.906	Bangor	ME
248	1.881	Parkersburg	WV	1.795	Southern NJ	NJ
249	1.750	Harrisonburg	MT	1.774	Pikeville	KY
min	1.282	Pikeville	KY	1.748	Wausau	WI

$C_{F,z}$ is the rate of Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and $C_{E,z}$ is the rate of Emotional-Gain crimes (Murder, Rape, and Aggravated Assault), in each Commuting Zone, z. Crime rates are expressed in Standardized Incidents per 100,000 people per year. A Standardized Incident is 1 Murder, 24 Robberies, 57 Motor Vehicle Thefts, or so on - see Table 2. t restricts by year (all 2010 here). Included are only the 250 largest Commuting Zones by population.

Table 6: Highest and Lowest Commuting Zones in 2010 Earnings and Inverse Earnings

	$Y_{t,z}$	CZ	State	$M_{t,z}$ (10^{-4})	CZ	State
max	\$44,336	Washington	DC	$\$^{-1}1.434$	Redding	CA
2	\$40,636	Northern NJ	NJ	$\$^{-1}1.362$	Medford	OR
3	\$40,384	Southern NJ	NJ	$\$^{-1}1.333$	Chico	CA
4	\$40,189	San Jose	CA	$\$^{-1}1.331$	Gallup	NM
5	\$39,953	San Francisco	CA	$\$^{-1}1.302$	Ocala	FL
6	\$39,216	Boston	MA	$\$^{-1}1.261$	Santa Rosa	CA
7	\$38,881	Bridgeport	CT	$\$^{-1}1.256$	Modesto	CA
8	\$38,715	Baltimore	MD	$\$^{-1}1.253$	Jackson	TN
9	\$38,686	New York City	NY	$\$^{-1}1.222$	Sumter	SC
10	\$37,862	Fredericksburg	VA	$\$^{-1}1.220$	Gastonia	NC
11	\$37,055	Minneapolis	MN	$\$^{-1}1.212$	Hot Springs	AR
12	\$36,469	Philadelphia	PA	$\$^{-1}1.206$	Bakersfield	CA
13	\$36,416	Seattle	WA	$\$^{-1}1.206$	Brownsville	TX
14	\$35,956	Denver	CO	$\$^{-1}1.193$	Pueblo	CO
15	\$35,735	Chicago	IL	$\$^{-1}1.192$	Santa Barbara	CA
60	\$31,582	Davenport	IA	$\$^{-1}1.072$	Rome	GA
61	\$31,581	Cleveland	OH	$\$^{-1}1.072$	Elkhart	IN
p75	\$31,556	Detroit	MI	$\$^{-1}1.071$	Jonesboro	AR
63	\$31,537	Little Rock	AR	$\$^{-1}1.070$	Tuscon	AZ
64	\$31,505	Peoria	IL	$\$^{-1}1.070$	Texarkana	TX
123	\$29,608	Fargo	ND	$\$^{-1}1.003$	Lexington	KY
124	\$29,597	Scranton	PA	$\$^{-1}0.999$	Las Vegas	NV
p50	\$29,587	Kalamazoo	MI	$\$^{-1}0.998$	Columbia	SC
126	\$29,504	Bellingham	WA	$\$^{-1}0.993$	Fort Smith	AR
127	\$29,335	Topeka	KS	$\$^{-1}0.992$	Charleston	SC
185	\$27,704	Paris	TX	$\$^{-1}0.935$	Jackson	MS
186	\$27,675	Tuscaloosa	AL	$\$^{-1}0.934$	Cincinatti	OH
p25	\$27,669	Corpus Christi	TX	$\$^{-1}0.933$	Austin	TX
188	\$27,667	Asheville	NC	$\$^{-1}0.930$	Racine	WI
189	\$27,660	Bloomington	IN	$\$^{-1}0.925$	Dallas	TX
246	\$24,797	Valdosta	GA	$\$^{-1}0.819$	Sioux Falls	SD
247	\$24,340	Sumter	SC	$\$^{-1}0.816$	Charleston	WV
248	\$23,921	Brownsville	TX	$\$^{-1}0.806$	Fredericksburg	VA
249	\$23,690	Hot Springs	AR	$\$^{-1}0.770$	Washington	DC
min	\$23,403	Laredo	TX	$\$^{-1}0.766$	Killeen	TX

Y_z is average post-tax earnings, and M_z is average post-tax inverse earnings, amongst age 18-64 people in the labor force in each Commuting Zone, z. t restricts by year (all 2010 here). Included are only the 250 largest Commuting Zones by population.

Table 7: Elasticities of the Crime Rate, $C_{g,t,z}$

		OLS		2SLS	
Parameter	Regressor	$g = E$	$g = F$	$g = E$	$g = F$
$\Theta_g = \vartheta_g/(1+\vartheta_g)$	$m_{t,z}$	0.355***	0.336***	**0.720*****	**0.783*****
		(0.038)	(0.035)	(0.080)	(0.075)
$\Theta_g\beta_g^y$	$y_{t,z}$	-0.008	0.137***	**0.069**	**0.383*****
		(0.044)	(0.041)	(0.072)	(0.067)
$\Theta_g\beta_g^q$	$q_{t,z}$	0.027***	0.049***	0.036***	0.048***
		(0.004)	(0.004)	(0.005)	(0.004)
$\Theta_g\beta_g^{Black}$	$p_{Black,t,z}$	0.100***	0.078***	0.110***	0.088***
		(0.006)	(0.006)	(0.007)	(0.006)
$\Theta_g\beta_g^{White}$	$p_{White,t,z}$	-0.098***	-0.079***	-0.023	0.019
		(0.029)	(0.027)	(0.033)	(0.031)
$\Theta_g = \vartheta_g/(1+\vartheta_g)$	$m_{t,z}$	0.391***	0.365***	**0.750*****	**0.788*****
		(0.039)	(0.037)	(0.079)	(0.075)
$\Theta_g\beta_g^y$	$y_{t,z}$	0.376***	0.432***	**0.394*****	**0.934*****
		(0.055)	(0.052)	(0.100)	(0.095)
$\Theta_g\beta_g^q$	$q_{t,z}$	0.026***	0.049***	0.028***	0.035***
		(0.004)	(0.004)	(0.005)	(0.004)
$\Theta_g\beta_g^{Black}$	$p_{Black,t,z}$	0.087***	0.065***	0.106***	0.077***
		(0.007)	(0.007)	(0.008)	(0.008)
$\Theta_g\beta_g^{White}$	$p_{White,t,z}$	0.076*	0.045	0.175***	0.255***
		(0.041)	(0.039)	(0.051)	(0.048)
$\Theta_g\beta_g^{Young}$	$p_{Young,t,z}$	-0.026	0.009	0.025	0.222***
		(0.065)	(0.062)	(0.073)	(0.069)
$\Theta_g\beta_g^{Jailed}$	$p_{Jailed,t,z}$	-0.019***	-0.009	-0.014**	-0.007
		(0.006)	(0.006)	(0.006)	(0.006)
$\Theta_g\beta_g^{College}$	$p_{College,t,z}$	-0.225***	-0.169***	-0.131***	-0.221***
		(0.021)	(0.020)	(0.031)	(0.030)
$\Theta_g\beta_g^{Married}$	$p_{Married,t,z}$	-0.721***	-0.553***	-0.577***	-0.646***
		(0.061)	(0.058)	(0.072)	(0.068)
$\Theta_g\beta_g^{Male}$	$p_{Male,t,z}$	0.697***	0.530***	0.326**	0.250**
		(0.115)	(0.110)	(0.131)	(0.124)
Observations		3,610	3,610	3,610	3,610
Weight		$N_{t,z}/N_t$	$N_{t,z}/N_t$	$N_{t,z}/N_t$	$N_{t,z}/N_t$

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the log crime rate. Observations are 722 Commuting Zones z, over 5 time periods t corresponding to {1980, 1990, 2000, 2010, 2016}. All columns include regional fixed effects, year effects, and year by region effects. N is population count. y is log average earnings. m is log average inverse earnings. q is log population density. p is log fraction of the local population. m, y, and p are restricted to non-institutionalized people of age [18, 64]; m and y are restricted further to those in the labor force. The 2SLS columns instrument for m and y with their Shift-Share counterparts. See previous tables for further definitions.

Table 8: Implied Parameter Values

	Baseline		Fully Controlled	
Parameter	g = E	g = F	g = E	g = F
Θ_g	0.720	0.783	0.750	0.788
$\Theta_g \beta_g^y$	0	0.383	0.394	0.934
$\beta_g^y = \Theta_g \beta_g^y / \Theta_g$	0	0.489	0.525	1.185
$\vartheta_g = \Theta_g/(1-\Theta_g)$	2.571	3.608	3.000	3.717

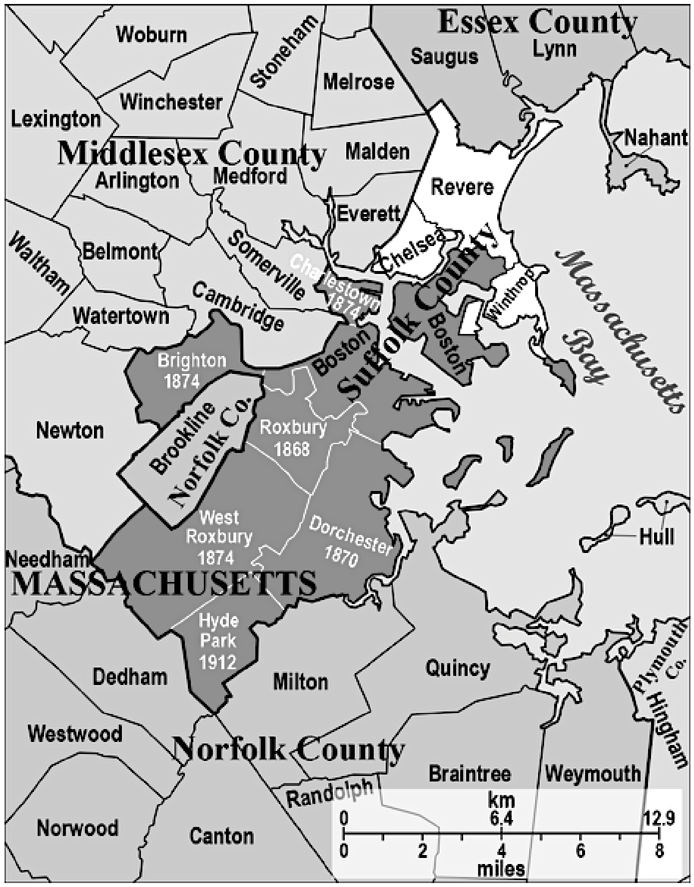
The first two rows are estimates taken from Table 7. The two lower rows calculate values of underlying structural parameters implied by the estimates in the first two rows.

Table 9: Counterfactual Responses

	$\check{\check{\vartheta}}_g = -0.2$		$\check{\check{\vartheta}}_g = +0.2$	
Response	g = E	g = F	g = E	g = F
$\check{\check{K}}_{g,z}$ *p90*	+0.091	+0.064	-0.092	-0.079
$\check{\check{K}}_{g,z}$ *p50*	+0.112	+0.092	-0.076	-0.062
$\check{\check{K}}_{g,z}$ *p10*	+0.140	+0.119	-0.062	-0.044
$\check{\check{C}}_{g,z}$ *p90*	-0.123	-0.107	+0.066	+0.046
$\check{\check{C}}_{g,z}$ *p50*	-0.101	-0.084	+0.082	+0.066
$\check{\check{C}}_{g,z}$ *p10*	-0.084	-0.060	+0.102	+0.085

$\check{\check{\vartheta}}_g$ is a counterfactual policy change. $\check{\check{\vartheta}}_g = -0.2$ would indicate that the policy change of choice is to decrease ϑ_g by 20%. The K and C responses give the resulting changes in expected punishment per crime and the crime rate, by Commuting Zone z. F indicates Financial-Gain crimes (Robbery, Burglary, Larceny, and Motor Vehicle Theft), and E indicates Emotional-Gain crimes (Murder, Rape, and Aggravated Assault). ϑ represents the government's distaste for criminal punishment. A lower value of ϑ coincides with a tougher stance on crime.

Figure 2: County Borders around Boston, MA



The map of the Boston, MA area illustrates why smaller geographical units such as cities or counties may be inappropriate to capture local crime markets. People frequently commute between Somerville, Cambridge, Boston (Basu and Ferreira, 2021, Florida and Mellander, 2016, Matarazzo et al., 2017), and Brookline. These four cities belong to three distinct counties - Middlesex, Suffolk, and Norfolk. Commuting Zones, defined based on actual commuting patterns in 1990, are meant to account for the vast majority of such local spillovers. The Boston CZ (see Figure 3) includes all of the counties pictured.

Figure 3: Boston, MA Commuting Zone

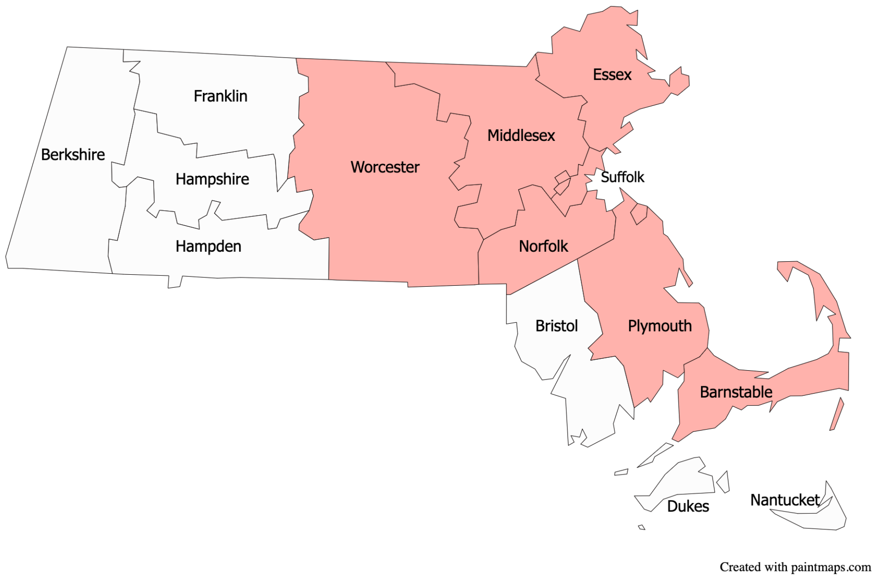
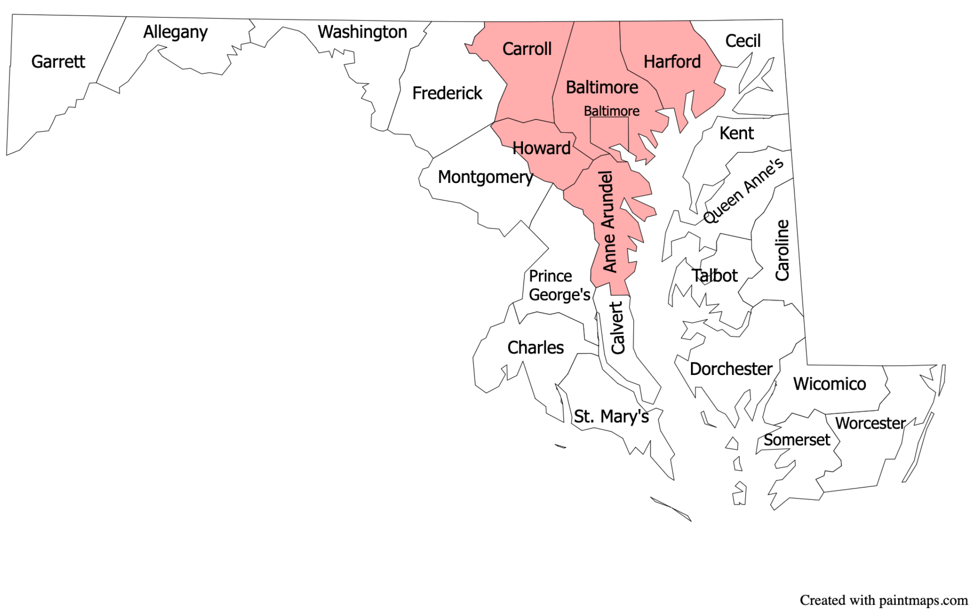


Figure 4: Baltimore, MD Commuting Zone



Most CZs are contained within state lines, but not all. The Washington, DC CZ contains Maryland counties from Frederick to St Mary's (above), as well as several Virginia counties.

A Appendix

A.1 Data Sources

As discussed in Section 3, we use county level summaries of the Decennial Census and American Community Survey (ACS), and also IPUMS USA microsamples of the same, in concert. We take county level summaries from Social Explorer,[28] particularly, 1970, 1980, 1990, and 2000 US Decennial Census summaries, and ACS 2008-2012 and 2015-2019 5-Year Estimate summaries. From IPUMS USA, we take the 1970 1% metro fm2, 1980 1% metro, 1990 1% unwt, 2000 1% unwt, 2010 ACS 1%, and 2019 ACS 1% samples. We use Dorn's crosswalk files[29] to integrate all of the above data into a panel of 722 Commuting Zones (CZs) by 6 decadal time periods.

At the CZ level, we compute a separate IPUMS measured and Social Explorer (SE) measured version of each variable, as available. In IPUMS, the source observations are of individual persons. Each variable we compute is a weighted average per CZ-year cell, weighted by perwt (person weight) and afactor, as described in Dorn (2009). In SE, the source observations are of counties, which group directly into CZs. For most SE measured variables, we take a population weighted average of the counties in each CZ-year cell.

For every variable at the CZ level, both in IPUMS and in SE, we also calculate versions of the same at two higher levels of aggregation, which we call Sub Region and Super Region. Sub Regions are the 9 regional divisions used by the Census,[30] but split into Urban and Rural CZs within each, for $9\times2 = 18$ in total. We define an Urban CZ as one in which, in the 1990 IPUMS sample, there were at least 80,000 people identified as living in an urban area as defined by the Census. Super Regions are the 4 main regions used by the Census, also split into Urban and Rural CZs within each, for 8 in total. The main regions are Northeast

[28]https://www.socialexplorer.com/explore-tables

[29]See Autor and Dorn (2013) and Autor, Dorn and Hanson (2013).

[30]The divisions are New England (CT, MA, ME, NH, RI, VT), Mid-Atlantic (NJ, NY, PA), East North Central (IL, IN, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DC, DE, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (AZ, CO, ID, MT, NV, NM, UT, WY), and Pacific (CA, OR, WA).

(New England, Mid-Atlantic), Midwest (East North Central, West North Central), South (South Atlantic, East South Central, West South Central), and West (Mountain, Pacific).

A.2 Data Transformations

We construct final versions of each variable at the CZ level based on weighted averages of several versions of the variable. The weights given to each version are to some extent increasing in the number of individual level observations underlying each. They include IPUMS sourced CZ level, Sub Region level, and Super Region level versions, and transformations of a closely related SE sourced variable at the CZ level, transformed at the Sub Region and Super Region levels. Transformations are of the form,

$$\mathrm{v}_{\mathrm{z}} = \frac{\mathrm{v}_{\mathrm{Z}}^{IPUMS}}{\mathrm{V}_{\mathrm{Z}}^{IPUMS}} \mathrm{V}_{\mathrm{z}}^{SE}$$

where V is a variable closely related to v that is available through SE, z is CZ, and Z is Region or Super Region. For example, with v as the average earnings of non-institutionalized people of age 18-64, V may be the average earnings of all people above age 15.

The use of these weighted averages serves three purposes. First, it guarantees a unique value to every CZ-year cell, for every variable. Second, it uses all available information in a consistent manner, distributing weight in accordance with the amount of information underlying each source estimate, which should improve overall accuracy. Third, the inclusion of Sub Region and Super Region level estimates within each final CZ level estimate accounts for some amount of spillover amongst neighboring CZs, albeit of an arbitrary magnitude. For the average CZ, the IPUMS CZ level estimate accounts for about 80% of the final estimate.

Our transformation technique, respective to any given case of v and V, hinges on the assumption that the ratio of v to V in z is equal to that in Z. v and V must be closely related for this assumption to be plausible. A concrete example may be: if age 16+ people in rural New Hampshire earn 4% more than the average for age 16+ people in rural New England, then age 18-64 people in rural New Hampshire are assumed to earn more 4% more than the average for age 18-64 people in rural New England. Symbolically,

$$v_z = \delta v_Z \ , \ V_z = \delta V_Z$$

where δ is a scalar, such as 1.04. Therefore,

$$\frac{v_z}{V_z} = \frac{v_Z}{V_Z}$$

We then assume that SE accurately measures V_z, and IPUMS accurately measures v_Z/V_Z.

A.3 Tax

As referenced in the model, we adjust earnings earnings by federal tax, redistributing half of the tax revenue as an in-kind earnings baseline $\underline{Y}_t$. The Census and ACS, via IPUMS USA, provide pre-tax earnings, Y_i, per individual person in the sample, i. National average pre-tax earnings are $\bar{Y}_t = N_t^{-1} \sum_i Y_i$, where N is population count, and t is time period. To arrive at post-tax earnings, we apply the tax function of Heathcote, Storesletten and Violante (2017),

$$\text{post-tax } Y_i \propto (\text{pre-tax } Y_i)^{1-\tau} \ , \ \tau = 0.181 \tag{16}$$

Heathcote et al. show that this function is a close match to the actual federal + state tax and transfer schedule in the US. That is, although the function is fit ($\tau = 0.181$) to capture state taxes (on average) in addition to federal tax, the function itself does not vary by state, so we refer to it throughout the paper simply as federal tax.

In conjunction with (16), it is necessary also to assume the aggregate amount of tax that is collected. We assume total tax is 20% of total pre-tax earnings, that is,

$$N_t^{-1} \sum_i (Y_i - \tilde{\tau} Y_i^{1-\tau}) = 20\% \ \bar{Y}_t \tag{17}$$

$Y_i - \tilde{\tau} Y_i^{1-\tau}$ being the net tax that is collected per individual i, and Y_i being pre-tax earnings.

(17) enables algebraic derivation of the constant of proportionality $\tilde{\tau}$, which resolves the exact value of each individual's tax payment $(Y_i - \tilde{\tau} Y_i^{1-\tau})$.

$$\tilde{\tau} = \frac{(1 - 0.2) \sum_i Y_i}{\sum_i Y_i^{1-0.181}} \tag{18}$$

Because $\tau = 0.181 > 0$, this is a progressive tax, taking more in percentage terms from individuals with higher pre-tax earnings Y_i.

Despite being a single function that is applied to all individuals i equally regardless of their location z, the tax (16) does affect the post-tax earnings averages Y_z across locations z differently. Table 6, which ranks CZs by their post-tax earnings averages Y_z, would not only list different values if not for the tax, but would be in a different order. For example, San Francisco may have a higher pre-tax average than Southern NJ, but with a relatively more unequal distribution. Because the tax is progressive, it takes a larger share from very high earning individuals, so may take substantially more in tax from San Francisco than it takes from Southern NJ. This could result in San Francisco's post-tax average falling slightly below Southern NJ's, despite the order being the other way around in pre-tax averages.

We assume that, in addition to paying the tax (16), people benefit from the tax revenue. Particularly, we assume that half of the tax budget is redistributed equally as in-kind benefits $\underline{Y}_t$. We add $\underline{Y}_t$ to each individual's post-tax earnings before calculating the averages $Y_{t,z}$ and $M_{t,z}$. Because the tax budget is 20% of total pre-tax earnings, $\underline{Y}_t$ is 10%, that is, $\underline{Y}_t = 10\% \ \bar{Y}_t$. This, in addition to the progressivity ($\tau > 0$) of the tax payment schedule, further moderates (reduces inequality in) the earnings distribution, both within locations z and across. That is, despite being paid out exactly equally across individuals, $\underline{Y}_t$ makes a larger difference in relative terms for people with lower earnings.

A.4 First and Second Order Conditions

This section shows that the equilibrium crime rate and expected level of punishment per crime given in Section 5 optimize the offender's utility function and government's loss function, respectively. Subscripts g (for crime type) and t (for time period) are suppressed for readability.

The offender i's utility function weighs his expected personal benefit per crime B_z against his expected punishment per crime K_z, in each Commuting Zone z.

$$\begin{aligned}
U(C_{i,z}) &= \frac{B_z}{1-\alpha}(C_{i,z})^{1-\alpha} - \frac{K_z}{2-\alpha}(C_{i,z})^{2-\alpha}(Y_{i,z} + \underline{Y})\ ,\ \alpha \in [0,1] \\
U'(C_{i,z}) &= B_z(C_{i,z})^{-\alpha} - K_z(C_{i,z})^{1-\alpha}(Y_{i,z} + \underline{Y}) \\
U''(C_{i,z}) &= -\alpha B_z(C_{i,z})^{-1-\alpha} - (1-\alpha)K_z(C_{i,z})^{-\alpha}(Y_{i,z} + \underline{Y})
\end{aligned} \tag{19}$$

The first order condition is:

$$U'(C_{i,z})\Big|\Big\{C_{i,z} = C^*_{i,z}\Big\} = 0$$

$$\Leftrightarrow\ B(C^*_{i,z})^{-\alpha} = K(C^*_{i,z})^{1-\alpha}(Y_{i,z} + \underline{Y})\ \Leftrightarrow\ C^*_{i,z} = (Y_{i,z} + \underline{Y})^{-1}\frac{B_z}{K_z}$$

The second order condition,

$$U''(C_{i,z})\Big|\Big\{C_{i,z} = C^*_{i,z}\Big\} < 0$$

follows straightforwardly from (19), regardless of $C^*_{i,z}$. Because $\alpha \in [0,1]$, at least one of $-\alpha$ and $-(1-\alpha)$ must be negative, and neither can be positive. C, B, K are all positive by construction. Therefore only negative terms are added together.

The government's loss function weighs the direct social cost of punishment K_z itself, against the social cost of crime, which responds to punishment indirectly, $C(K_z)$.

$$\begin{aligned}
L(K_z) &= \vartheta^{-1}K_z^{\vartheta} + C(K_z) = \vartheta^{-1}K_z^{\vartheta} + \frac{M_z B_z}{K_z} \\
L'(K_z) &= K_z^{\vartheta-1} - \frac{M_z B_z}{K_z^2} \\
L''(K_z) &= (\vartheta-1)K_z^{\vartheta-2} + 2\frac{M_z B_z}{K_z^3}
\end{aligned} \tag{20}$$

The first order condition is:

$$\mathrm{L}'(\mathrm{K_z})\bigg|\bigg\{\mathrm{K_z} = \mathrm{K_z^*}\bigg\} = 0$$

$$\Leftrightarrow\ (\mathrm{K_z^*})^{\vartheta-1} = \frac{\mathrm{M_zB_z}}{(\mathrm{K_z^*})^2}\ \Leftrightarrow\ (\mathrm{K_z^*})^{\vartheta-1+2} = \mathrm{M_zB_z}\ \Leftrightarrow\ \mathrm{K_z^*} = (\mathrm{M_zB_z})^{1/(1+\vartheta)}$$

Because $\mathrm{L}(\mathrm{K_z})$ is a loss function, it is optimized at a minimum, where the second derivative is positive. Thus the second order condition is:

$$\mathrm{L}''(\mathrm{K_z})\bigg|\bigg\{\mathrm{K_z} = \mathrm{K_z^*} = (\mathrm{M_zB_z})^{1/(1+\vartheta)}\bigg\} > 0$$

The derived $\mathrm{K_z^*}$ can be plugged in to (20), yielding:

$$\mathrm{L}''(\mathrm{K_z})\bigg|\bigg\{\mathrm{K_z} = \mathrm{K_z^*}\bigg\} = (\vartheta - 1)(\mathrm{M_zB_z})^{(\vartheta-2)/(1+\vartheta)} + 2(\mathrm{M_zB_z})^{1-3/(1+\vartheta)}$$

Notice that $1-3/(1+\vartheta) = (1+\vartheta-3)/(1+\vartheta) = (\vartheta-2)/(1+\vartheta)$. Therefore the above simplifies to:

$$(\vartheta - 1 + 2)(\mathrm{M_zB_z})^{(\vartheta-2)/(1+\vartheta)} = (1 + \vartheta)(\mathrm{M_zB_z})^{(\vartheta-2)/(1+\vartheta)} > 0$$

$(1 + \vartheta)$ must be positive because $\vartheta > 0$, and $\mathrm{M_zB_z}$ are positive by construction.

A.5 Counterfactual Responses

The main step in deriving counterfactual responses to changes in ϑ involves inverting the equilibrium equations (10) and (11). With (crime type) g and (time) t subscripts suppressed, these equilibrium equations are:

$$\mathrm{K_z} = (\mathrm{M_zB_z})^{1/(1+\vartheta)} = (\mathrm{C_z})^{1/\vartheta}$$

$$\mathrm{C_z} = (\mathrm{M_zB_z})^{\vartheta/(1+\vartheta)} = (\mathrm{K_z})^{\vartheta}$$

Inverted versions of the same are:

$$\begin{aligned} \mathrm{M_zB_z} &= (\mathrm{K_z})^{1+\vartheta} \\ \mathrm{M_zB_z} &= (\mathrm{C_z})^{(1+\vartheta)/\vartheta} \end{aligned} \tag{21}$$

Because nothing in $\mathrm{M_z}$ or $\mathrm{B_z}$ is a function of ϑ, counterfactual equilibrium values $\check{\mathrm{K}}_\mathrm{z}$ and $\check{\mathrm{C}}_\mathrm{z}$ follow from the counterfactual $\check{\vartheta}$ of choice as:

$$\begin{aligned} \check{\mathrm{K}}_\mathrm{z} &= (\mathrm{M_zB_z})^{1/(1+\check{\vartheta})} \\ \check{\mathrm{C}}_\mathrm{z} &= (\mathrm{M_zB_z})^{\check{\vartheta}/(1+\check{\vartheta})} \end{aligned}$$

Although $\mathrm{M_zB_z}$ is not directly observed, it can be replaced using the inverted factual equilibrium (21):

$$\begin{aligned} \check{\mathrm{K}}_\mathrm{z} &= ((\mathrm{K_z})^{1+\vartheta})^{1/(1+\check{\vartheta})} = (\mathrm{K_z})^{(1+\vartheta)/(1+\check{\vartheta})} \\ \check{\mathrm{C}}_\mathrm{z} &= ((\mathrm{C_z})^{(1+\vartheta)/\vartheta})^{\check{\vartheta}/(1+\check{\vartheta})} = (\mathrm{C_z})^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})} \end{aligned} \tag{22}$$

Of primary interest are how each variable changes in the counterfactual relative to its factual counterpart:

$$\check{\check{\vartheta}} := (\check{\vartheta} - \vartheta)/\vartheta \ , \ \check{\check{\mathrm{K}}}_\mathrm{z} := (\check{\mathrm{K}}_\mathrm{z} - \mathrm{K_z})/\mathrm{K_z} \ , \ \check{\check{\mathrm{C}}}_\mathrm{z} := (\check{\mathrm{C}}_\mathrm{z} - \mathrm{C_z})/\mathrm{C_z} \tag{23}$$

For example, $\check{\check{\vartheta}} = 0.1$ would mean that the counterfactual policy choice is to increase ϑ by 10%. Combining (22) and (23):

$$\begin{aligned} \check{\check{\mathrm{K}}}_\mathrm{z} &= ((\mathrm{K_z})^{(1+\vartheta)/(1+\check{\vartheta})} - \mathrm{K_z})/\mathrm{K_z} = (\mathrm{K_z})^{(1+\vartheta)/(1+\check{\vartheta})-1} - 1 = (\mathrm{K_z})^{-\vartheta\check{\check{\vartheta}}/(1+\check{\vartheta})} - 1 \\ \check{\check{\mathrm{C}}}_\mathrm{z} &= ((\mathrm{C_z})^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})} - \mathrm{C_z})/\mathrm{C_z} = (\mathrm{C_z})^{(\check{\vartheta}+\vartheta\check{\vartheta})/(\vartheta+\vartheta\check{\vartheta})-1} - 1 = (\mathrm{C_z})^{\check{\check{\vartheta}}/(1+\check{\vartheta})} - 1 \end{aligned}$$

To simplify further - to functions of only the observed $\mathrm{C_z}$, estimated ϑ, and chosen $\check{\check{\vartheta}}$ - apply the facts that $\mathrm{K_z} = \mathrm{C_z}^{1/\vartheta}$ (equilibrium condition), and $\check{\vartheta} = \vartheta(1 + \check{\check{\vartheta}})$ (definition of $\check{\check{\vartheta}}$):

$$\check{\check{\mathrm{K}}}_\mathrm{z} = (\mathrm{C_z})^{-\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$
$$\check{\check{\mathrm{C}}}_\mathrm{z} = (\mathrm{C_z})^{\check{\check{\vartheta}}/(1+\vartheta(1+\check{\check{\vartheta}}))} - 1$$

The above are equilibrium formulas that predict how expected punishment per crime K and crime rates C would adjust under a different social loss function. A higher value of ϑ (so $\check{\check{\vartheta}} > 0$) would mean society (the government) placing increased weight on the pain inherent in criminal punishment in and of itself. A lower value of ϑ (so $\check{\check{\vartheta}} < 0$) would mean the opposite - a society that is less sensitive to the pain of punishment (relative to the harm caused by crime), and therefore a tougher stance against crime.