Uncovering the Gender Participation Gap in the Crime Market

Evelina Gavrilova
Nadia Campaniello

DISCUSSION PAPER SERIES

I Z A

Uncovering the Gender Participation Gap in the Crime Market

Evelina Gavrilova
Norwegian School of Economics

Nadia Campaniello
University of Essex
and IZA

ABSTRACT

Uncovering the Gender Participation Gap in the Crime Market[1]

There is little research on the gender variation in the crime market. We document a gender gap in criminal activities, based on property crimes, using data from the U.S. National Incident Based Reporting System from 1995 to 2015. We show that there is a gender participation gap, with only 30 percent of the crimes being committed by females. We try to explain the gender participation gap by focusing on incentives to commit crime, such as criminal earnings and probability of arrest. We show that on average females earn 13 percent less than males while they face a 9 percent lower likelihood of arrest. We find that males respond more to changes in illegal earnings, with an elasticity of 0.36, while females are less responsive with an elasticity of 0.23. Both genders respond equally to changes in the probability of arrest, with an elasticity around -0.14. Using a Blinder- Oaxaca type decomposition technique, we find that differences in incentives explain about 8 percent of the gender participation gap, while differences in responsiveness to changes in incentives, especially illegal earnings, explain about 56 percent of the gap. The fact that females behave differently than males has implications for the heterogeneity in response to crime control policies.

Corresponding author:

Nadia Campaniello
University of Essex
Department of Economics
Wivenhoe Park
Colchester, Essex CO4 3SQ
UK
E-mail: ncampa@essex.ac.uk

---

[1] This is a revised version posted in March 2018. The original version circulated in April 2015 included fewer years in the empirical analysis and omitted important fixed effects.

1　Introduction

In the past years there has been considerable attention given to the gender gaps in wages and employment. Yet, almost no notice has been given to the gender gap in an illegal setting and none on the gender gap in criminal earnings. Most research in the economics of crime has focused on male perpetrators (Levitt & Miles, 2007; Freeman, 1999), with the implicit assumptions that female crime is so little that it is of no consequences or that policy implications have external validity across genders. Yet, female crime has been growing in the last decades, as we show with the percentage share of women who are incarcerated in Figure 1, and it is not possible anymore to assume it away. One of the main reason for which there is a little research on crime is the data limitation on demographic details about the criminal perpetrators.

In this article we want to fill the "gap" in research on the gender crime gap. In the economic model of crime criminal participation is intrinsically connected with the possible illegal earnings and detection probability. We use detailed matched offense-arrest data on property (including white-collar crimes) from the the U.S. *National Incident Based Reporting System* (NIBRS) for the period 1995-2015. These data offer the unique opportunity to link the participation decision to the likelihood of arrest and the illegal earnings. We match the NIBRS data to population statistics and compute cohort specific crime rates[2], where cohorts are defined based on interval of age, race, gender and county allowing us to explore the gender participation decisions in crime in light of the associated illegal earnings and arrest probabilities.

We start off by quantifying the gaps in crime for participation, illegal earnings and arrest probability. Only around 30 percent of the crimes are committed by women. We find that females earn 7 percent less than males and face a similar probability of arrest. When we explore the heterogeneity with respect to crime type we find that females participate most in shoplifting. In the absence of this crime, we observe a more severe earnings gap of 13 percent and a 9 percent lower probability of arrest for females. We interpret this as possible evidence that females sort into shoplifting as a way to obtain higher criminal earnings, yet they suffer from the higher detection and arrest rate associated with this particular crime.

In our second set of results we try to relate the participation rates to earnings and arrest probabilities. First, we assume that criminals rationally form their expectations on the criminal incentives[3] based past experience. This mechanism is similar to the one described in Lochner (2007). Then we use these predicted values to see how female and male crime rates respond to changes in expected illegal earnings and in the probability of arrest, allowing us to identify gender-specific elasticities.

We find strong evidence that past incentives predicted future ones, generating a strong "first stage." As for the responsiveness, we find that males have a higher elasticity of illegal earnings of 0.36 compared to 0.23 for females. Both genders have a similar elasticity of arrest, predicting that one percent increase in the probability of arrest leads to a 0.14 percent decrease in crimes committed. By exploiting a partial Blinder-Oaxaca decomposition we try to measure how much of the participation decision is due to different criminal incentives and different elasticities. We find that if females were more "manly" with respect to incentives

---

[2]Henceforth, we will refer to these crime rates as participation rates.

[3]In the standard economic model of crime both higher earnings and a lower arrest probability can persuade a criminal to commit a crime. Therefore, we will refer to these two variables as criminal incentives. The probability of arrest deters criminals, generating a negative incentive rather than a positive one. In our model, we allow males and females to respond differently to these incentives.

and responsiveness to them this would reduce the participation gap by 64 percent.

The main implication of our findings concerns the external validity of previous studies that have been conducted only on a sample of male offenders. We show that males and females respond differently to incentives.

Crucially, our elasticities with respect to the probability of arrest could be compared to previous literature. In a recent review of the literature, Chalfin & McCrary (2017) show that estimated elasticities with respect to an increase in the police manpower are in the range of -10 to -20 percent. Our estimates fall well within this range and are significant at conventional levels.

To the best of our knowledge, there are no studies with US data with estimated elasticities with respect to illegal earnings to which we can compare our estimates. Our elasticities are roughly in the range 25-40 percent as estimated by Draca *et al.* (2015) on UK data. Our estimates imply that policies that aim to reduce the illegal earnings of criminals would impact males more than females.

Our research is of interest to policy makers that want to decrease crime. We provide evidence for the heterogeneity in response between genders to policies that alter the incentives to engage in crime. More specifically, if the policy maker wants to discourage males to participate in crime, this would be most effective with policies that decrease the value of potential earnings. Examples for such policies are black market regulation, where pawn shops could be placed under additional surveillance or additional security for higher value items (see D'Este, 2014).

Related Literature. We contribute to several strands in the literature. Most prominently, in the Handbook of Labour Economics, Freeman (1999) acknowledges the gap in studies about the gender variation in crime and underlines that there are no studies by economists that analyze the large differences in the participation of males and females. Since then there has been scant response to this apparent gap and we are one of the first to fully investigate this research question.

On one hand, there are a few papers that look at the criminal market in isolation. The earliest economic study on female criminals, Bartel (1979), investigates the determinants of female participation in crime through an Ehrlich type model of time division. The author finds that probabilities of conviction and arrest have a deterrent effect on females in some property crimes. Our results are in line with her findings. Gavrilova (2017) investigates the incentives for assortative matching in a criminal pair and finds that females are likely discriminated against when matching with males.

On the other hand, there is a small literature that tries to identify the participation decision by looking at shifts in the opportunity cost of crime. Recently, Corman *et al.* (2014) find that the 1996 "welfare reform" in the U.S., aimed at incentivizing female work, led to a decrease in female arrests for serious property crimes by $4.4 - 4.9\%$. Cano-Urbina *et al.* (2016) look at the effect of education on female crime. They find that one more year of school for females reduces, on average, property and violent crimes by 50%, while they do not find any effect on white collar crime. They argue that the effect of education on crime for females is probably due to changes in marital opportunities rather than labor market opportunities. In our paper we control for employment opportunities and wage that represent the opportunity costs of being involved in criminal activities.

A recent study by Beatton *et al.* (2017) examines the convergence of the gender crime gap in Australia. They find gaps that are in line with the ones we find in the US context. They show that over the course of the last 20 years these gaps have contracted by 30 percent. In descriptive evidence here we show that over almost the same period of time the US crime

gap has contracted by 28 percent, again a very similar number. Overall, female participation in crime seems to be increasing around the world.

By focusing our analysis on the illegal earnings of criminals we are contributing to the understanding of "the most understudied element of crime" (Draca & Machin, 2015). Recent literature has only attempted to approximate the illegal earnings of criminals, such as Draca *et al.* (2015), while we have more precise information on the value of the property stolen.

Finally, we use NIBRS data on property offenses that include white collar crimes, a quite neglected and understudied typology of crimes in the literature.

2 Descriptives

2.1 Data

For our analysis we use the *National Incident Based Reporting System* (NIBRS). This dataset contains records on the universe of crime incidents for a given year for a given law-enforcement agency in the United States. The data are not representative for the United States as a whole, as many agencies do not submit reports and the expansion of data collection is on-going. A typical observation is a coded report about a criminal incident. It contains the number of perpetrators, their demographic characteristics and crime codes. We match the report data to the arrest data, in order to see whether a perpetrator was arrested. We observe both arrested and non-arrested criminals. Criminal earnings are recorded regardless of whether there was an arrest.

We limit our analysis to property crimes (including white collar crimes). We do this for two reasons. First, property crimes are more common than other crimes. The violent crime rate over the period 2016 is 386 crimes per 100,000 inhabitants, while the property crime rate is 2,451 crimes. White collar crimes themselves seem to be even more prevalent, according to survey evidence from Huff *et al.* (2010). Second, for property and white collar crimes we can observe a relevant incentive such as the illegal earnings, while for violent crimes the incentives are difficult to quantify.

In Table 1 we detail the types of crimes that we analyze. Once we select these crime incidents we have 45 million observations on criminals over the period 1995 to 2015. We select individuals between 15-44 years of age, of black or white race. This selection is mainly guided by the availability of control variables, as we want to approximate the opportunity cost of crime in the best way we can. In addition, we exclude the two other races Asian and Native American that are provided in the data, because they comprise a too small fraction (3 percent each in the dataset).

In order to understand the criminal participation decision we define "pseudo-individuals" and construct a synthetic panel (see Deaton, 1985). We aggregate crimes within "cohorts" and divide by the corresponding population, thus getting a measure of crime rates and variation in the participation decision. The resulting unbalanced panel is treated as pseudo-individuals that can be tracked over time. The cohorts are defined based on: interval of age (15-24, 25-34 and 35-44[4]), race (black and white), gender (male, female), and county. The panel spans over a period of 20 years.

We take data on population (by age, gender, race, year, and county) from the Wide-ranging Online Data for Epidemiologic Research (WONDER). Data on average wages and

---

[4]Data on the general population at county level are just available until the age of 44 and we know that 75 percent of the crimes are committed by perpetrators from the selected age groups.

employment rates (by age, gender, race, decade, and state) are taken from the CENSUS Integrated Public Use Microdata Series (IPUMS-USA).

2.2 Descriptive Evidence

Table 2 shows the summary statistics by gender. Across cohorts we observe that there are on average 20 females committing crimes versus 39 males. These numbers correspond to crime rates of 713 per 100,000 inhabitants for males and 331 per 100,000 inhabitants for females. Both of these crime rates are below conventional levels of property crime rates in the US, as cited above for 2016 the crime rate was 2,451. Likely the difference comes from the fact that in our sample we keep offenders whose demographic characteristics were well observed, in a certain age group, while in the official statistics reports are aggregated for all types of offenders, both unknown and observed. Females have a smaller likelihood of arrest of 35 percent versus 36 percent for males. We also show that females have less earnings, at $1724 versus $1928 for males.

In Figure 2 we show how these statistics have evolved over time. We observe that the female/male ratio in criminal participation increased over the 20 year sample period. Concurrently with this increase in female participation, we observe that women, on average, have decreased their criminal earnings and slightly increased their probability of arrest with respect to men.

In Figure 3 we present the gender participation gap as a function of the earnings gap (top row) and the arrest gap (bottom row) for the different crime types. In panel (a) we show that when females' relative earnings increase, their participation in criminal activities decreases but this relationship seems to be driven by shoplifting. Panel (b) shows that the relationship becomes positive, as one would expect if criminals respond to incentives, when shoplifting is excluded. In panel (c) we show that females participate more in criminal activities when their relative likelihood of arrest is lower, which is again coherent with economic theory. A shown in panel (d), shoplifting seems to be in line with the overall trend.

Since behind the averages in Figure 2 there might be significant hidden heterogeneity, in Figure 4 we compare the densities of earnings and arrest probability for males and females. We present earnings on the top row and arrest probability on the bottom row. The first column shows the overall data, while the second column excludes shoplifting. In panel (a) we observe that female earnings are bimodal, while male ones are more to the right in location. In panel(b) when excluding shoplifting the two distributions are more similar and comparable. However, the male distribution still remains to the right of the female, mirroring the averages from Table 2 and showing that males earn more than females. In panel (c) we observe the density of the probability of arrest for males and females and we notice the sharp spike in the right part of the density of females. In panel (d) we show that this spike is driven by shoplifting, and as we exclude it we again arrive at more comparable distributions. The fact that shoplifting can drive the spike in panel (c) of Figure 4 could be due to the crime being reported only if an arrest takes place. Since this would bias the elasticity of crime with respect to the probability of arrest, we exclude shoplifting in the main analysis (but include it in the robustness checks).

3 Model of Crime

3.1 Theory

Following Becker (1968), we assume that an individual compares the expected utility of committing a crime with the expected utility of not committing a crime. The expected benefits from a crime are the illegal earnings. The expected cost is the sanction length. Both benefits and costs hinge on the realization of the probability of detection and arrest. The expected utility of not committing the crime is based on the opportunity cost of crime: being engaged in the legal labor market. We introduce in the model a gender-specific component, in order to test for difference between genders in the subsequent sections of this paper.

Therefore, an individual from gender g, $g \in \{m, f\}$, decides to be involved in a criminal activity if a gender-specific function $f_g()$ of costs and benefits is larger than an individual idiosyncratic error Υ_g. Υ_g can measure any unobserved determinants of crime such as sociocultural and family factors. By allowing the function to be gender-specific we allow each gender to respond differently to their own expectations:

$$f_g(E(Y_g), E(P_g), E(W_g), E(L_g), E(J_g)) > \Upsilon_g \ , \tag{1}$$

We take the expectation of the criminal over variables like illegal earnings (denoted by Y), probability of arrest (P), legal wage (W), employment (L) and sanctions (J). In order to arrive to a estimable equation, we log-linearize the function $f()$ with respect to the variables in the model and note that when we sum over all the cases, in which $f_g() > \Upsilon_g$, delivers the crime rate C_g:

$$c_g = \beta_g^1 y_g + \beta_g^2 p_g + \beta_g^3 w_g + \beta_g^4 l_g + \beta_g^5 j_g + \varepsilon_g \tag{2}$$

where lower case letter denote logged quantities. Assuming a specific functional form for $f(\cdot)$ allows us to gauge some intuition about the importance of the different pieces. If criminals have constant relative risk aversion utility functions, the different factors that enter the utility function are unrelated to each other, and criminals can either work or engage in crime, condition (1) becomes:[5]

$$f_g(\cdot) = \hat{Y}_g^{1-\rho_g}/(1-\rho_g) - \hat{P}_g \hat{J}_g - (\hat{W}_g \hat{L}_g)^{1-\rho_g}/(1-\rho_g) > \Upsilon_g. \tag{3}$$

Log-linearizing and assuming that Υ_g is uniformly distributed it is easy to show that the elasticity of crime with respect to criminal earnings is proportional to $(1 - \rho_g)$, while the elasticity of crime with respect to the probability of arrest would be the same for males and females. This would imply that more risk loving criminals are, the more they respond to changes in the expected illegal earnings.

Using a log-log specification the coefficients measure elasticities. Equation 2 implies that a gender crime participation gap (c_m vs c_f) could be both due to (1) gender-specific differences in key variables such as illegal earnings y_g and (2) over the way criminals respond to such incentives, for e.g. β_g^1. The subscript g on the coefficients β_g denotes that the effect of the variable could be different by gender. With the subscript g on the variables we model the fact that the expectation could be formed over the gender-specific expected values for the variables. That is, the expectations for males are based on the variables like earnings and arrest of other males.[6]

---

[5] "Hats" indicate criminals' perceptions.

[6] We use the variable g in subscript here in order to describe better the data in the next subsection when

More formally, we explicitly model the expectations of criminals as a function of the previous realization of the variables. Lochner (2007) shows with survey data that indeed criminals form expectations on the probability of arrest by observing realizations in their own peer group. Similarly, we assume that the expectation over the probability of arrest p_g and earnings y_g is formed by observing the peer group within the cohort of criminals. We define cohorts at the crime type-, gender-, race-, age group- and location level, so that the cohort captures the peer group and this peer group is significantly homogeneous. This is similar to Lochner (2007), who also defines a narrow peer group over which the expectations of criminals are modeled.

3.2   Estimation

In this section we present our empirical methodology. There are two main reasons for the presence of a participation gap. First, there could be a difference in the incentives that each gender faces such as probability of arrest or earnings.[7] Second, even if the incentives were the same, maybe the two genders react differently to them. In the next subsection, we show how we want to describe the differences between the two genders in terms of illegal earnings and arrest - two of the main incentives in the crime participation decision. Second, we show how sensitive this participation decision is to changes in incentives of probability of arrest and illegal earnings.

3.2.1   Incentives

In Section 2 and Figure 3 we document significant differences between males and females in terms of illegal earnings, arrest and participation. To describe how earnings and arrest vary with gender and other factors we estimate the specifications of the following type for the criminal earnings and probability of arrest:

$$z_{git} = \beta D(g = f)_{git} + X'_{git}\gamma + \varepsilon_{git} \tag{4}$$

where z_{git} is the dependent variable (either the log illegal earnings or the log arrest rate) for criminals of gender g, in cohort i and year t. $D(g = f)$ is an indicator function equal to one when the gender is female. A $\beta = 0$ would imply that the dependent variable does not vary with the gender. X is a vector containing personal-cohort traits like race, age, average wage and unemployment rate.

We include offense fixed effects in order to account for the effects of different offenses on the unconditional gap. For example, we expect that a criminal would earn more in auto theft crimes than in shoplifting and if males specialize in the former, while females specialize in the latter, this would earn a high unconditional gap. In order to control for county specific heterogeneity such as police presence in any given year, we include county-year fixed effects. Furthermore, in some specifications we also includes state-offense fixed effects to control for differences in sanctions across different states. Finally, we cluster the standard errors at the county level, in order to account for correlation of residuals over time within county.

---

we define the estimating equations.

[7]The probability of arrest is a deterrent to crime rather than an "economic" incentive. In the interest of the following exposition, we will refer to both earnings and arrest probability as incentives, while keeping in mind that "incentives" is heterogeneous category.

3.2.2 Responsiveness to Incentives

Once we map the log differences between the two genders in criminal earnings and arrest probabilities, we turn to the participation decision. Starting from Equation 2 we obtain:

$$c_{git} = \beta_g^1 y_{git} + \beta_g^2 p_{git} + X_{git}' \delta_g + \varepsilon_{git}, \text{ for } g \in \{m, f\}, \tag{5}$$

where c_{git} is the log crime rate defined as the number of crimes committed by criminals in gender g in a given year t by people within a cohort i. The cohort is defined by age group, race, and county, and the crime rate is divided by the general population in the same cohort. y is the log of expected illegal earnings, p is the log of the expected probability of arrest. X is a vector containing sociodemographic variables, specific to the cohort, such as wage, employment rate, race and age group.

Similar to Equation 4, in some specifications we include county-year fixed effects in order to control for county specific heterogeneity, such as police presence, in any given year. In some specifications we include state-offense fixed effects to control for differences in criminal sanctions across different states. Finally, we cluster the standard errors at the county level. Note that we estimate equation 5 separately for males and females.

There are two potential problems when using the contemporaneous values y_{git} and p_{git}: i) reverse causality due to the potential simultaneity between the incentives and the decision to commit a crime (for example, crime congestion might lower the likelihood of apprehension) and ii) due to the yearly aggregation criminals' expectations might be based on future crimes, introducing additional measurement error.

Little is known about how criminals form their economic expectations about the criminal incentives but a few articles have developed theories that we use to set up our empirical methodology. As mentioned in Lochner (2007) *"Theories developed in Sah (1991) and Lochner (2004) stress that the probability of arrest is learned from others or through one's own experiences...Beliefs about the probability of arrest are likely to depend on an individuals own (past) criminal behavior and arrest outcomes, the criminal and arrest outcomes of others around him, and more general signals that may come from local arrest rates or neighborhood conditions."*

In other words, current expectations are a function of past realizations. We assume that past realizations influence current realizations only through the expectations of criminals, once we account for all state-offense-specific and county-year-specific influences. Individuals who have been operating in the criminal markets in the past or who know someone who has been operating in the past, use past experience to form expectations about the future. Moreover, if expectations are rational then average realizations at time t should be equal to the ex-ante expectations (*t-1*), and these should be a function of all the information available to criminals up until time *t-1*, including previous realizations.

In modeling the expectations we use a two step procedure, where in the first stage we obtain $\widehat{p_{git}}$ and $\widehat{y_{git}}$ and in the second step these measures are used in Equation 5 to determine whether males and females have different responses to the predicted values of incentives. These equations are an instrumental variable 2SLS regression, where $y_{gi,t-1}$ and $p_{gi,t-1}$ are used as instruments for y_{git} and p_{git}.[8]

---

[8]Given that the first stage equation is a dynamic specification, it implies that the coefficient α_g will suffer from a downward bias (Nickell, 1981). This implies that the estimated $\widehat{\alpha_g}$ will be smaller than the true α_g, if $\alpha_g > 0$. Therefore, the first stage F-statistic will be lower than the "true" one and the instrument will be stronger than we can observe. In addition, the identified effects in the second stage will be smaller than what we observe. This bias is alleviated by the length of the panel, which is 20 years in our case. Nonetheless, we

Measurement error. Illegal earnings and probability of arrest are likely to be impacted differently by a measurement error. The probability of arrest suffers less from measurement error because we have precise information on those who get arrested. For illegal earnings, measurement error after averaging might come from victim reports or law-enforcement policies thus biasing our results towards zero. This would be an issue in our IV specifications if such measurement error was correlated across years. Given the small likelihood of victimization it is unlikely for a victim to report crimes in consecutive years and so the serial correlation would have to come from group behaviour (i.e. biased reporting by gender, race, typology of crime, etc.). Adding the corresponding fixed effects can flexibly account for these sources of measurement error. As for police policies, in our empirical specification we include interacted year-county fixed effects which can account for year to year changes in recording crimes by the police. Moreover, measurement error might be an issue if i) there is a systematic gender bias in collecting data on illegal earnings (we believe this to be implausible, especially for property crimes where the victims have an incentive to report the crime and the value of the stolen items) and ii) males and females select into different crimes that have different measurement error. While this could easily be going on, since we have a log-specification and we control for the typology of crimes committed, any systematic bias or differential measurement error in reporting would be absorbed by the fixed effects.

3.2.3 Blinder-Oaxaca decomposition

In order to determine the importance of the criminal incentives, like earnings y and probability of arrest p, on the gender crime gap, we use a partial Blinder-Oaxaca decomposition, limited to y and p. We are interested in how much of the difference between male crime participation, measured by c_{mit}, and female crime participation, c_{fit}, can be explained by incentives (endowments in the jargon of the decomposition) and responsiveness (the betas). To do this we define two counterfactuals, using the estimates obtained from Equation 5.

In the first counterfactual we will assign to females the male incentives, thereby changing females' "endowment". In this way we want to observe how the difference in earnings and arrest probability between the two genders contributes to the observed participation gap. Defining $\widehat{c_{fit}} = y_{mit}\widehat{\beta}_f^1 + p_{fit}\widehat{\beta}_f^2 + Z'_{fit}\widehat{\gamma}_f$ to be the predicted female crime rate based on Eq. 5, we derive the counterfactual in the following way:

$$c_{fit}^m = y_{mit}\widehat{\beta}_f^1 + p_{mit}\widehat{\beta}_f^2 + Z'_{fit}\widehat{\gamma}_f = \widehat{c_{fit}} + (y_{mit} - y_{fit})\widehat{\beta}_f^1 + (p_{mit} - p_{fit})\widehat{\beta}_f^2. \tag{6}$$

In other words, substituting the female endowments, y_{fit} and p_{fit}, with the male ones, y_{mit} and p_{mit}, generates the counterfactual female crime rate. The new crime rate c_{fit}^m can be interpreted as the crime rate in a world where females have the same earning and the same probability of arrest as males. Figure 5 shows the cumulative distribution functions of the counterfactual crime rate, the male crime rate and the female one. In this figure we want to observe how the distance between $\widehat{c_{mit}}$ and c_{fit}^m compares to the initial distance between $\widehat{c_{mit}}$ and $\widehat{c_{fit}}$. The difference in distances would show how much of the initial gap arises due to differences in the possible earnings and arrest probabilities for each gender.

To determine how much of the observed gap is due to differences in the estimated responses to earnings and arrest (the βs), we define the second counterfactual:

$$c_{fit}^\beta = \widehat{c_{fit}} + (\widehat{\beta}_m^1 - \widehat{\beta}_f^1)(y_{fit} - \underline{y}_f) + (\widehat{\beta}_m^2 - \widehat{\beta}_f^2)(p_{fit} - \underline{p}_f) \tag{7}$$

---

comment on its effect in results in Section 4.

This counterfactual corresponds to a world where females have the same elasticities as males have, but keep their initial endowments. In this expression we adjust the female coefficients β_f in the estimation of the predicted female crime rate, by replacing them with their male counterparts.[9]

To determine how much of the observed gap arises due to differences in the estimated elasticities, we again plot the cumulative distribution functions in Figure 5. In the second panel of the figure we again compare the difference between the new predicted female crime rate c_{fit}^{β} and the male crime rate $\widehat{c_{mit}}$ to the actual gap in crime participation, between $\widehat{c_{mit}}$ and $\widehat{c_{fit}}$. This comparison shows how much of the crime participation gap is due to differences in the way males and females respond to incentives.

4 Results

4.1 Differences in incentives

In Table 3 we present results for the illegal earnings gap. As explained in Section 2.2, shoplifting is a special crime, therefore we exclude it from the first four columns and include it in the last column as a robustness check. In column 1 we show females have 13 percent lower criminal earnings. In the next three columns we add progressively more fixed effects, but the gap remains stable at around 13 percent. Our preferred specification is that of column 4 with all the controls and the fixed effects. By adding these fixed effects we want to account for jurisdiction specific policing responses, for the sanction for a specific crime in each state, as well as availability of criminal targets related to the business cycle. In column 5 we select a subsample of daylight crimes committed between 8 AM and 7 PM. We do this robustness check in order to alleviate concerns on reporting bias: it might be the case that if the female perpetrators were not well observed they would be reported as male. During daylight he gap is still negative and is even larger in magnitude. Females earn 15 percent less than males in daylight crimes, which is similar to the previous estimates and shows that reporting biases are unlikely to drive the gap.

Finally, in column 6 of Table 3 we include shoplifting as a robustness checks. We find that the gap diminishes to 7 percent, meaning that the earnings gap in shoplifting is in favor of females and is significant enough to attenuate the estimates on the earnings gap in all crimes.

In Table 4 we present the results for the arrest gap. In the first column the arrest gap is negative at 5 percent, meaning that females are 5 percent less likely to be arrested than males. However, the gap increases to 9 percent as we control for more heterogeneity by adding different fixed effects. As before, our preferred estimate is from the specification in column 4, where we observe an arrest gap of 9 percent. The arrest gap increases in daylight crimes to 15 percent, suggesting a reporting bias where females could have been reported as males. The increase in the relative number of new "females" would drive down the relative probability of arrest and result in a smaller gap as in the previous columns, conditional on there being no other significant differences in crime participation within the 24 hours of the day. We assume that in the daylight hours most perpetrators are clearly observed.

In the last column of Table 4 we show that shoplifting drives the estimates on the average gap in arrest rates, which now decreases to 0 percent. This reveals that shoplifting is a crime in which females face a higher likelihood of arrest than males.

---

[9]Subtracting the minimum values $\underline{y}^f$ and $\underline{p}^f$ normalizes the intercepts and rotates the counterfactual around its minimum value (rather than around the initial intercept, where $y = 0$ and $p = 0$).

Overall, conditional on crime participation we find that females earn 7 to 13 percent less than males and face a lower likelihood of being arrested.

4.2 Differences in Responsiveness

Table 5 shows our estimates for the first stage and reduced form using alternative specifications with and without county fixed effects and their interaction with year fixed effects, and with and without state fixed effects and their interaction with typology of offense fixed effects.

In the top panel we observe that the lag of the (log) probability of arrest and the lag of (log) illegal earnings are good predictors of, respectively, the (log) probability of arrest and (log) illegal earnings. The F-statistics is, in all the specifications, well above conventional levels at 187 for Males in our preferred specification and 94 for Females. In the bottom panel of Table 5 it is reassuring that the coefficients have the right sign. An increase in the illegal earnings is associated with an increase in the crime rate and an increase in the probability of arrest is associated with a decrease in the crime rates.

In Table 6 we present the 2SLS estimates on the gender-specific response elasticities. In each additional column we control for more sources of heterogeneity such as police presence, sentence length and business cycle effects by adding fixed effects and their interactions. We show in the odd columns with the header "Male" that a 100 percent increase in the expected illegal earnings would lead to a 36 percent higher participation of males in crime. Similarly, the elasticity of expected criminal earnings for females in our preferred specification in the last column is 23 percent. On average, we observe that females have always a significantly lower elasticity of illegal earnings than males. As mentioned in Section 3.1, this could be driven by females being more risk averse than males.[10]

With respect to the probability of arrest, we find that for both males and females, the elasticity is between -14 and -22 percent. In our preferred specification in columns 7 and 8, the male and female coefficients are not significantly different and are around -14 percent. Males and females respond similarly to an increase in the probability of arrest.

Correcting for the Nickell bias,[11] we find somewhat larger first stage coefficients and we end up with an elasticity of earnings for males of around 0.26 and an elasticity of earnings for females around 0.15. The elasticities of arrest decrease to 13 percent. All elasticities remain significant at conventional levels.

With respect to the control variables, we find that being black is associated with an increase in the crime rate. Offenders aged from 15-24 contribute more to the crime rate than offenders aged 25-34 compared to the excluded category 35-44. The crime rate is also associated with the average wage and employment rate in all specifications, even though the coefficients change in the different specifications that we use. These results seem to be highly dependent on the inclusion of the fixed effects. In our preferred specification in column 7 and 8, an increase in the wage is associated with an increase in the crime rate, whereas the employment rate is negatively associated with the crime rate.

To sum up, we find that both males and females respond to incentives. They are both more responsive to changes in the illegal earnings and remain equally deterred by increases in the probability of arrest. Males tend to respond more than females.

---

[10]For an overview of experimental evidence that women appear to be more risk averse than men see Eckel & Grossman (2008).

[11]We do that by applying the equation (19) from Nickell (1981) which approximates the bias to the estimated coefficient. Then we take the ratio between the reduced form coefficient and the corrected first stage to arrive at the bias corrected coefficients, which are reported in Table 8.

In Table 7, we show OLS estimates. We find that the estimates of the elasticities are smaller, which is consistent with the simultaneity issues outlined before.

Robustness Checks. In Table 8 we perform robustness checks with the aim to see whether our estimates depend on the particular specification we used. First of all, in the first set of column we report our estimates from Table 6 corrected for the Nickell bias, which were commented above. Second, to be sure that our results are not biased by the smaller dimension of the cohorts in the sample, we cut the sample at the median population and take the cohorts with larger populations. Almost all elasticities increase in magnitude. In larger cohorts, an increase in the earnings of 100 percent would lead to an increase in participation by 42 percent for males and 35 percent for females, which is larger than the average response. At the same time, the response to the probability of arrest increases for males to -18 percent and slightly decreases to -12 percent for females. Overall, when we use larger cohorts criminals seem to be more responsive to incentives.

In the last panel of Table 8 we add shoplifting to the sample. We find larger earnings elasticities and smaller arrest elasticities for both genders. We interpret this as evidence that shoplifting is a crime that is recorded only when arrest takes place, which biases the corresponding coefficients towards zero (as already shown in Figure 4).

4.3 Blinder-Oaxaca decomposition

In Figure 5 we present the cumulative density functions for the Blinder-Oaxaca decomposition exercise. In Panel (a) we plot the result from equation 6, where we assign to females the male endowments in terms of earnings and probability of arrest. In Panel (b) we plot the result from equation 7 where we assign to females the same responsiveness to criminal incentives as males. In both panels we also plot the predicted cumulative distribution functions of male and female crime rates. The distance between them exemplifies the participation gaps that we observe in Section 2.2.

In Panel (a) we show that if females had male endowments, the participation gap would change by only 8 percent. This is shown by the vicinity between CDF of female crimes and the counterfactual predicted from equation 6. This is driven by the fact that in Section 4.2 we show that both gender react similarly to increases in the probability of arrest. The large gap that remains is partly due to the fact that males have a larger elasticity of illegal income than females. Here it is important to highlight that one has to take a stand on how to treat the crime fixed effects. Keeping the female fixed effects we are changing the endowments without allowing females to switch across crimes.

In Panel (b) we show that if females had the same elasticities as males, then the gap would be smaller by 56 percent. Given that the two genders respond similarly to increases in the probability of arrest, this shift is driven by an increase in the elasticity of illegal earnings. This result implies that differences in risk aversion may go a long way in explaining differences in criminal attitudes.

Together, differences in endowments and in the elasticities explain 64 percent of the gap.

5 Conclusion

We motivate our research with the fact that little is known about female criminals and about what may be driving the gender crime gap.

Over the last 50 years there has been an increase in convergence between the roles of males and females in the legal market, yet there is little evidence on the existence or development

of gaps in the criminal market.

From a historical perspective, in the 70s concurrently with the women emancipation movements, there have been concerns about an increase in the female participation in crime (Simon, 1976). In the last years the broad social context has been redefining sex roles. Weisheit (1984) cites a dominant hypothesis that female participation in crime would increase as social sex roles converge. In line with these theories, in the present study, that focusses on the United States over the period 1995-2015, we show evidence of the presence of a gender gap in property crimes that has been shrinking over time.

In our data we observe that in 1995 there was one female criminal for every 3 male ones, in 2015 this ratio increased to one female for every 2 male criminals.

In order to be able to understand such convergence it is important to determine why there is a gender gap in the first place. In our paper we unpack the gender gap in criminal participation in three steps. First, we quantify the observed difference in illegal earnings and probability of arrest (the so called "criminal incentives" in the Becker's model of crime) across gender controlling for several of characteristics and for a number of fixed effects and their interactions. The mapping of these differences in criminal incentives across gender is our first contribution. We observe that, *ceteris paribus*, women earn 13 percent less than men when committing a crime but they face a 9 percent lower probability of arrest. There are a host of different explanations for this: ability, choices, effort, search costs as well as underlying risk aversion.

We also observe that these differences between male and female criminals shrink over time, which may contribute to a reduction in the participation gap if criminals were responding to such incentives. And this is what we try to establish in the second part of the study. We model the participation decision and provide evidence that males and females increase their criminal engagement when the expected illegal earnings go up and when the probability of arrest goes down. We also find that females and males respond differently to changes in illegal earnings, which might signal underlying differences in preferences, for example a higher risk aversion among women.

Finally, in order to measure how such differences can contribute to the gender gap we use a Blinder-Oaxaca decomposition. Differences in criminal incentives explain 8 percent of the gender gap, while differences in the elasticities explain a much larger part of the gap, 56 percent. The remaining 36 percent of the gap could be driven by differences in the disutility of prison (Mastrobuoni & Rivers, forthcoming), differences in discount factors (Mastrobuoni & Rivers, 2016), differences in the opportunity cost of crime, due for example, to child rearing, etc.

More elaborate differences may also play a role. Drawing from the economic analysis of socially constructed identities (Akerlof & Kranton, 2000), it is easy to make the conjecture that stereotypes can play a role in the participation decision. If crime was a masculine job then entering females would de-value the masculinity image and may thus be ostracized by their male counterpart. This would lead to females being less likely to be initiated into crime than their male counterparts. On the flip side, if females think that crime is non-feminine, they would be less likely to participate. This role of stereotypes coincides with the findings of Steffensmeier (1980), who notes that males are less likely to choose a female partner because they consider females to be less trustworthy and more governed by passions. Gavrilova (2017) shows that there is a bias against females in criminal group formation. If females do not get initiated into crime in the same way as males, then maybe they participate less. If this would be the case, then the illegal market offers a setting in which discrimination is a blessing in disguise. A blessing, because policy makers and society at large likely prefer lower crime

participation rates. Disentangling the importance of all these mechanisms is left for future research, as it requires a great deal of information about criminals that is currently unavailable.

6   Acknowledgements

References

National Incident-Based Reporting System. United States Department of Justice. Federal Bureau of Investigation, 1995-2010. *Inter-university Consortium for Political and Social Research.*

Akerlof, George A, & Kranton, Rachel E. 2000. Economics and Identity.

Bartel, A.P. 1979. Women and Crime: An Economic Analysis. *Economic Inquiry*, **17**(1), 29–51.

Beatton, Tony, Kidd, Michael P, & Machin, Stephen. 2017. *Gender Crime Convergence over Twenty Years.* Tech. rept.

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

Cano-Urbina, Javier, Lochner, Lance, *et al.* . 2016. *The Effect of Education and School Quality on Female Crime.* Tech. rept. University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Chalfin, Aaron, & McCrary, Justin. 2017. Criminal deterrence: A review of the literature. *Journal of Economic Literature*, **55**(1), 5–48.

Corman, Hope, Dave, Dhaval M, & Reichman, Nancy E. 2014. Effects of welfare reform on women's crime. *International Review of Law and Economics*, **40**, 1–14.

Deaton, Angus. 1985. Panel Data From Time Series of Cross Sections. *Journal of Econometrics*, **30**(1-2), 109–126.

D'Este, Rocco. 2014. *The Effect of Stolen Goods Markets on Crime: Evidence from a Quasi-Natural Experiment.* Tech. rept. University of Warwick, Department of Economics.

Draca, Mirko, & Machin, Stephen. 2015. Crime and Economic Incentives. *Annual Review of Economics*, **7**(1), 389–408.

Draca, Mirko, Koutmeridis, Theodore, & Machin, Stephen J. 2015. The changing returns to crime: do criminals respond to prices?

Eckel, Catherine C, & Grossman, Philip J. 2008. Men, Women and Risk Aversion: Experimental Evidence. *In: Handbook of Experimental Economic Results.* Citeseer.

Freeman, Richard. 1999. The Economics of Crime. *Handbook of Labor Economics, Vol 3c.*

Gavrilova, Evelina. 2017. A Partner in Crime: Assortative Matching and Bias in the Crime Market. *Journal of Economic Behavior and Organization.* Forthcoming.

Huff, Rodney, Desilets, Christian, & Kane, John. 2010. National public survey on white-collar crime. *Retrieved October*, **28**, 2012.

Levitt, Steven D, & Miles, Thomas J. 2007. Empirical Study of Criminal Punishment. *Handbook of Law and Economics*, **1**, 455–495.

Lochner, Lance. 2004. *Belief Updating and Criminal Careers.*

Lochner, Lance. 2007. Individual perceptions of the criminal justice system. *American Economic Review*, **97**(1), 444–460.

Mastrobuoni, Giovanni, & Rivers, David. 2016. Criminal discount factors and deterrence.

Mastrobuoni, Giovanni, & Rivers, David A. forthcoming. Optimizing Criminal Behavior and the Disutility of Prison. *The Economic Journal.*

Nickell, Stephen. 1981. Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417–1426.

Sah, Raaj K. 1991. Social osmosis and patterns of crime. *Journal of political Economy*, **99**(6), 1272–1295.

Simon, Rita J. 1976. American Women and Crime. *The Annals of the American Academy of Political and Social Science*, **423**(1), 31–46.

Steffensmeier, Darrell J. 1980. Sex Differences in Patterns of Adult Crime, 1965-77: A Review and Assessment. *Social Forces*, 1080–1108.

Weisheit, Ralph A. 1984. Women and Crime: Issues and Perspectives. *Sex roles*, **11**(7-8), 567–581.

Figure 1: Fraction of Female Prisoners and Female Workers in the U.S

Notes: On the left axis we plot the fraction of females prisoners while on the right axis we plot the fraction of women in the labour force in the US over the period 1930-2009. Sources: National Archive of Criminal Justice Data. Online at https://www.icpsr.umich.edu/icpsrweb/NACJD; U.S. Census Bureau, Statistical Abstract of the United States: 2011, Law Enforcement, Courts and Prisons.

Figure 2: Differences between Males and Females
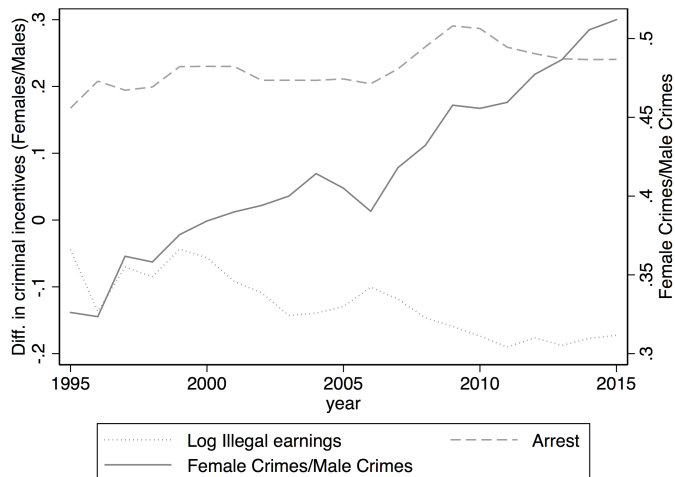
Notes: In this graph we plot the evolution of criminal participation, arrest and illegal earnings of the female to males ratio over the period 1995-2015. On the left axis we plot illegal earnings and arrest while on the right axis we plot the fraction of female criminals.

Figure 3: Gender Participation Gap in Crime and Incentives
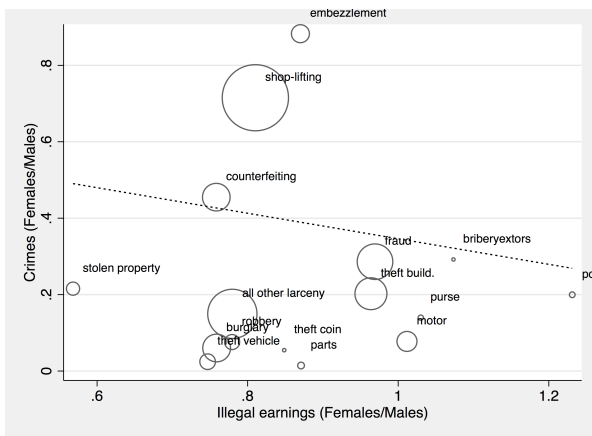
Panel (a)
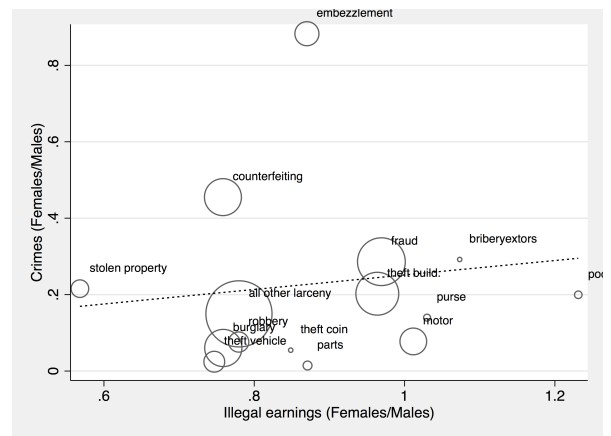
Panel (b)
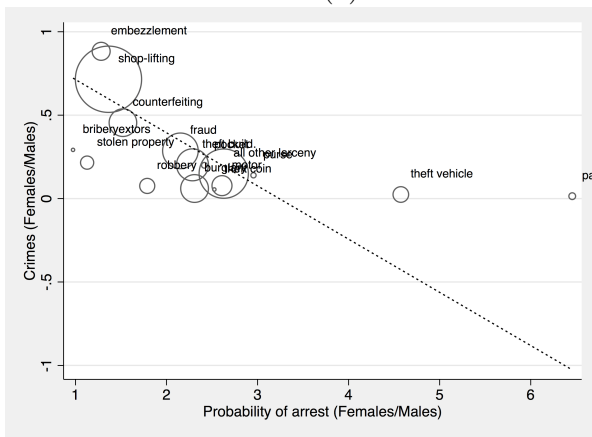
Panel (c)
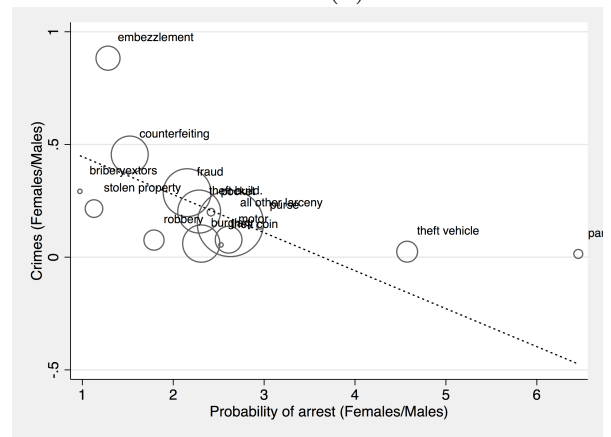
Panel (d)

Notes: The Y-axis denotes the ratio of female to male criminals. Values above 1 denote crimes with a higer female than male participation. Values below 1 indicate crimes with a higher male participation. The interpretation on the X-axis is similar. Circles are proportional to the number of females who commit each crime. Figures on the left of the panel include shoplifting while those on the right do not. The size of the circle is proportional to the fraction of females in each crime.

Figure 4: Density of Illegal Earnings and Arrest for Males and Females

Panel (a)

Panel (b)
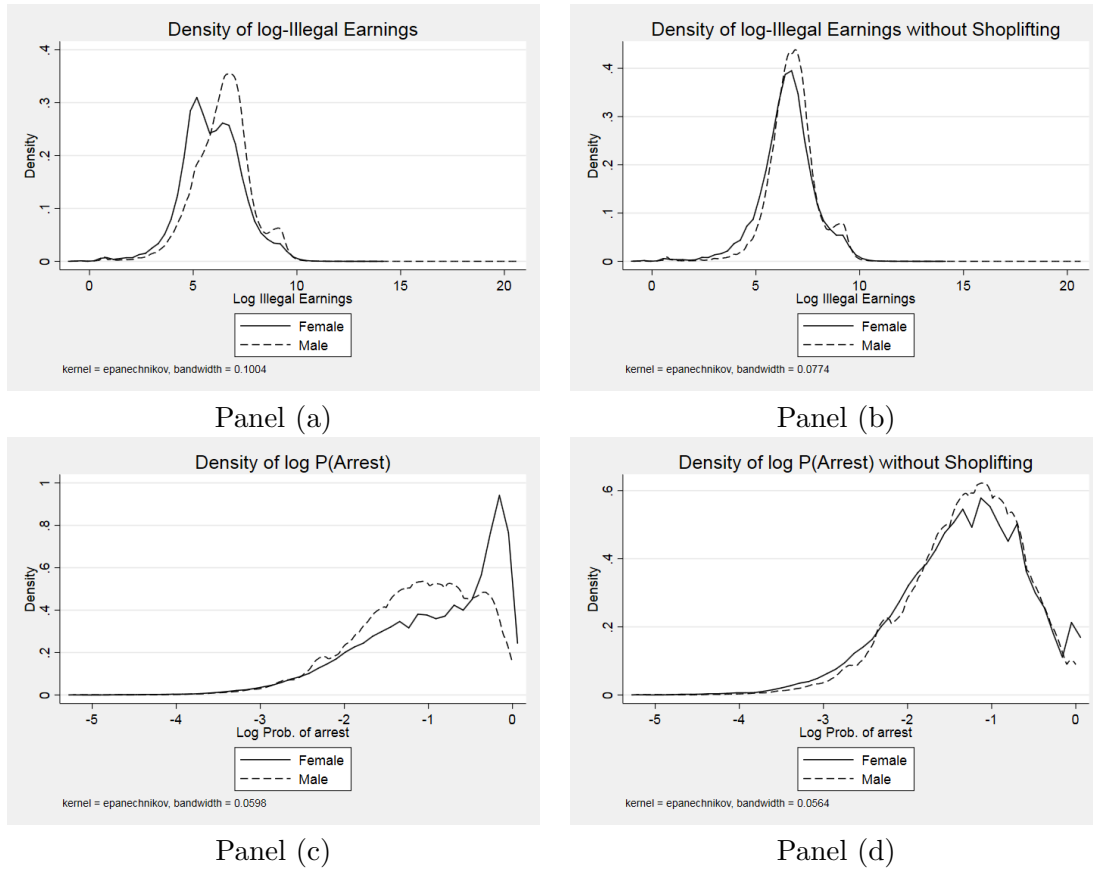
Panel (c)

Panel (d)

Notes: Here we plot the kernel density of log of illegal earnings and arrests for males and females. Figures on the left of the panel include shoplifting while those on the right do not.

Figure 5: Crime Gaps in the Counterfactual Scenarios
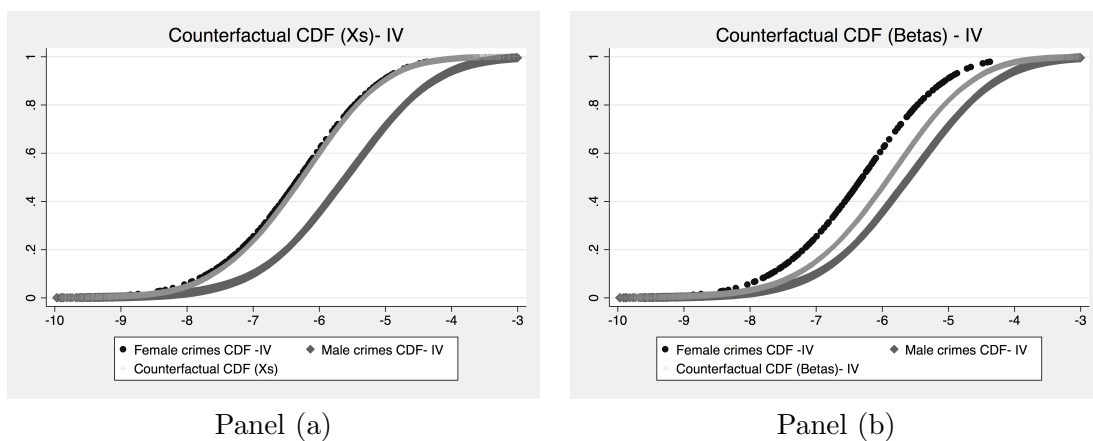
Panel (a)

Panel (b)

Notes: In panel (a) we present the male, female and female counterfactual (with the male endowments) cumulative density function of crime rates using estimates from the IV setup. The females counterfactual CDF reduces the gap between the two CDFs. In panel (b) we present the male, female and female counterfactual scenario (with the male coefficients) cumulative density function of crime rates using estimates from the IV setup. The females counterfactual CDF reduces the gap between the two CDFs.

Table 1: Crime Codes

UCR Code	Description	Subsummed into category*
120	Robbery	
210	Extortion/Blackmail	Bribery and Extortion
220	Burglary	
231	Pocket-picking	
232	Purse-snatching	
233	Shoplifting	
234	Theft from Building	
235	Theft from Coin-Operated Machine	
236	Theft from Motor Vehicle	
237	Parts	
238	All Other Larceny	
240	Motor Vehicle Theft	
250	Counterfeiting/Forgery	
261	Swindle	Fraud
262	Credit Card ATM Fraud	Fraud
263	Impersonation	Fraud
264	Welfare Fraud	Fraud
265	Wire Fraud	Fraud
270	Embezzlement	
280	Stolen Property Offenses	
510	Bribery	Bribery and Extorsion

*We sum over 261 Swindle, 262 Credit Card ATM Fraud, 263 Impersonation, 264 Welfare Fraud, 265 Wire Fraud and call the new variable "Fraud". We sum over 210 Extortion/Blackmail and 510 Bribery and call them "Bribery and Extorsion"

Table 2: Summary Statistics for the Synthetic Panel (1995-2015)

	Males				Females			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Crime variables:								
Number of crimes	38.99	88.23	2	5359	20.12	25.75	2	473
Crime rates (x 100,000)	713.22	1120.99	0.922	55445.55	331.11	532.74	0.869	26562.5
Illegal Earnings	1928.97	5635.30	0.4	1434448	1724.91	4566.78	0.4	434066.50
Prob. of arrest	0.36	0.20	0.01	0.99	0.35	0.19	0.01	0.98
Demographic variables:								
Age 15 - 24	0.42	0.49	0	1	0.40	0.49	0	1
Age 25 - 34	0.34	0.47	0	1	0.35	0.48	0	1
Age 35 - 44	0.24	0.43	0	1	0.25	0.43	0	1
Black	0.28	0.45	0	1	0.26	0.44	0	1
White	0.72	0.45	0	1	0.74	0.44	0	1
Population	11858.18	18314.76	6	265977	13847.97	18697.15	45	244929
Wage and salary income	21645.77	14615.84	940	83930.83	14825.88	8340.59	1392.86	37647.86
Employment rate	0.67	0.18	0	1	0.64	0.11	0.25	1
Crime type indicator variables:								
Embezzlement	0.04	0.19	0	1	0.06	0.24	0	1
Fraud	0.12	0.33	0	1	0.16	0.37	0	1
Counterfeiting	0.08	0.27	0	1	0.11	0.32	0	1
Bribery and extorsion	0.00	0.01	0	1	0.00	0.00	0	0
Pocket-picking	0.00	0.04	0	1	0.00	0.04	0	1
Purse-snatching	0.00	0.05	0	1	0.00	0.03	0	1
Theft from building	0.11	0.32	0	1	0.12	0.33	0	1
Theft from coin operated machine	0.00	0.03	0	1	0.00	0.01	0	1
Theft from or of motor-vehicle and its parts	0.07	0.25	0	1	0.03	0.18	0	1
Parts	0.01	0.11	0	1	0.00	0.06	0	1
All other larceny	0.23	0.42	0	1	0.25	0.43	0	1
Motor vehicle theft	0.09	0.29	0	1	0.07	0.25	0	1
Burglary	0.16	0.37	0	1	0.12	0.32	0	1
Robbery	0.05	0.22	0	1	0.04	0.19	0	1
Stolen property offenses	0.04	0.19	0	1	0.02	0.15	0	1
Observations		150,782				90,394		
Shoplifting	0.15	0.36	0	1	0.20	0.40	0	1
Observations with Shoplifting		177,644				114,281		

Table 3: Differences in Illegal Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
			Log Illegal Earnings			
Female	-0.133***	-0.147***	-0.127***	-0.128***	-0.152***	-0.0730***
	(0.0160)	(0.0138)	(0.0108)	(0.00979)	-0.0115	(0.0102)
Black	0.122***	0.0343	0.00743	-0.00662	-0.0380**	0.0854***
	(0.0296)	(0.0375)	(0.0208)	(0.0174)	-0.0183	(0.0153)
Age 15 - 24	-0.0304	-0.00282	-0.136***	-0.125***	-0.158***	-0.0791**
	(0.0562)	(0.0340)	(0.0331)	(0.0290)	-0.0349	(0.0310)
Age 25 - 34	0.00994	0.00703	-0.0132	-0.0105	-0.0236**	0.0241**
	(0.0154)	(0.0127)	(0.00967)	(0.00908)	-0.0116	(0.00976)
Log Wage and salary income	0.428***	0.358***	0.249***	0.239***	0.199***	0.187***
	(0.0564)	(0.0404)	(0.0355)	(0.0325)	-0.0365	(0.0340)
Employment rate	-0.920***	-0.550***	-0.457***	-0.370***	-0.201*	-0.0139
	(0.222)	(0.147)	(0.116)	(0.112)	-0.121	(0.129)
Log Population	0.221***	0.136***	0.134***	0.120***	0.109***	0.0967***
	(0.0117)	(0.0237)	(0.0160)	(0.00946)	-0.0106	(0.00908)
Year FE	yes	yes	yes	yes	yes	yes
County FE	no	yes	yes	yes	yes	yes
Year*County FE	no	no	yes	yes	yes	yes
Offense FE	no	no	yes	yes	yes	yes
State*Offense FE	no	no	no	yes	yes	yes
Sample		No shoplifting			Daylight and no shoplifting	All
Observations	241,176	241,176	241,176	241,176	166,213	291,925
R-squared	0.103	0.170	0.532	0.559	0.497	0.702

Notes: The dependent variable in all regressions is the log of illegal earnings. Excluded categories are: *white* for the race, people in the *age between 35 and 44* for the age group and *year 1995* for the year dummies. Estimation includes fixed effects where noted with "yes". Estimates are weighted by the size of the cohort. Standard errors clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Differences in Probability of Arrest

	(1)	(2)	(3)	(4)	(5)	(6)
			Log Probability of arrest			
Female	-0.0532**	-0.100***	-0.0894***	-0.0909***	-0.0385***	0.00169
	(0.0229)	(0.0228)	(0.0129)	(0.0124)	(0.0128)	(0.00847)
Black	-0.0574	-0.0274	-0.0255	-0.0336	0.0142	-0.0681***
	(0.0488)	(0.0275)	(0.0226)	(0.0204)	(0.0234)	(0.0170)
Age 15 - 24	0.636***	0.384***	0.293***	0.284***	0.254***	0.233***
	(0.0944)	(0.0826)	(0.0414)	(0.0386)	(0.0415)	(0.0289)
Age 25 - 34	-0.00928	-0.0434*	-0.0635***	-0.0616***	-0.0591***	-0.0692***
	(0.0306)	(0.0230)	(0.0209)	(0.0207)	(0.0211)	(0.0184)
Log Wage and salary income	0.206**	0.0475	-0.0204	-0.0135	-0.0369	0.0127
	(0.102)	(0.0762)	(0.0468)	(0.0448)	(0.0482)	(0.0381)
Employment rate	0.648*	0.592***	0.653***	0.578***	0.573***	0.369**
	(0.364)	(0.207)	(0.174)	(0.165)	(0.177)	(0.143)
Log Population	-0.0527***	-0.0229	-0.0338**	-0.0258*	-0.0124	-0.0123
	(0.0201)	(0.0218)	(0.0154)	(0.0146)	(0.0173)	(0.0113)
Year FE	yes	yes	yes	yes	yes	yes
County FE	no	yes	yes	yes	yes	yes
Year*County FE	no	no	yes	yes	yes	yes
Offense FE	no	no	yes	yes	yes	yes
State*Offense FE	no	no	no	yes	yes	yes
Sample			No shoplifting		Daylight and no shoplifting	All
Observations	241,176	241,176	241,176	241,176	166,213	291,925
R-squared	0.076	0.336	0.558	0.605	0.566	0.706

Notes: The dependent variable in all regressions is the log of the probability of arrest. Excluded categories are: *white* for the race, people in the *age between 35 and 44* for the age group and *year 1995* for the year dummies. Estimation includes fixed effects where noted with "yes". Estimates are weighted by the size of the cohort. Standard errors clustered at the county level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: First Stage and Reduced Form

	(1) Males	(2) Females	(3) Males	(4) Females	(5) Males	(6) Females	(7) Males	(8) Females
				Panel A: First stage				
				Log Illegal Earnings				
Log Lag Illegal Earnings	0.267***	0.194***	0.163***	0.119***	0.142***	0.113***	0.115***	0.0866***
	(0.0178)	(0.0105)	(0.00994)	(0.00789)	(0.00669)	(0.00702)	(0.00598)	(0.00629)
Log Lag Prob. of arrest	-0.0558***	-0.121***	-0.0325***	-0.0345***	-0.0381***	-0.0353***	-0.0270***	-0.0347***
	(0.0127)	(0.0127)	(0.00910)	(0.00981)	(0.00814)	(0.00965)	(0.00735)	(0.00952)
				Log Probability of arrest				
Log Lag Prob. of arrest	0.498***	0.455***	0.314***	0.272***	0.291***	0.257***	0.242***	0.208***
	(0.0112)	(0.0139)	(0.0103)	(0.0106)	(0.00920)	(0.00942)	(0.00914)	(0.0100)
Log Lag Illegal Earnings	-0.00249	-0.0112***	-0.000835	-0.000578	-0.00192	-0.00163	-0.000362	-0.00182
	(0.00231)	(0.00231)	(0.00148)	(0.00183)	(0.00145)	(0.00179)	(0.00136)	(0.00172)
First stage F-stat	108.7	163.1	133.2	113.2	226.7	129.1	186.9	93.94
				Panel B: Reduced form				
				Log Crime rates				
Log Lag Illegal Earnings	0.000485	-0.0336***	0.0518***	0.0308***	0.0513***	0.0304***	0.0417***	0.0204***
	(0.00719)	(0.00694)	(0.00320)	(0.00320)	(0.00352)	(0.00336)	(0.00257)	(0.00250)
Log Lag Probability of arrest	-0.0854***	-0.0513*	-0.0777***	-0.0693***	-0.0628***	-0.0546***	-0.0445***	-0.0366***
	(0.0240)	(0.0274)	(0.0106)	(0.0107)	(0.0105)	(0.0108)	(0.00969)	(0.00885)
Controls								
Offense FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	no	no	yes	yes	yes	yes	yes	yes
Year*County FE	no	no	no	no	yes	yes	yes	yes
State*Offense FE	no	no	no	no	no	no	yes	yes
Test of equality between males and females (p-value)								
Log Illegal earnings		0.000		0.000		0.000		0.000
Log Probability of arrest		0.000		0.000		0.000		0.222
Observations	150,782	90,394	150,782	90,394	150,782	90,394	150,782	90,394

Notes: The dependent variable is on top of the estimates column. Control variables include race (*black*), age groups (*age 15-24* and *age 25-34*), *log wage and salary income*, and *employment rate*. Estimation includes fixed effects where noted with "yes". In the reduced form we test whether males and females have coefficients on the variables of interest that are significantly different from zero. Standard errors clustered at the county level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Log Crime rates				
	Male	Female	Male	Female	Male	Female	Male	Female
Log Illegal Earnings	0.000222	-0.182***	0.317***	0.259***	0.359***	0.266***	0.361***	0.233***
	(0.0271)	(0.0379)	(0.0261)	(0.0295)	(0.0252)	(0.0321)	(0.0246)	(0.0325)
Log Prob. of arrest	-0.171***	-0.161***	-0.214***	-0.222***	-0.169***	-0.176***	-0.143***	-0.137***
	(0.0473)	(0.0611)	(0.0324)	(0.0382)	(0.0354)	(0.0425)	(0.0380)	(0.0438)
Black	1.385***	0.670***	1.412***	1.035***	1.383***	1.001***	1.367***	0.970***
	(0.0684)	(0.0468)	(0.0454)	(0.0397)	(0.0465)	(0.0388)	(0.0454)	(0.0381)
Age 15 - 24	0.215	-0.404**	1.085***	0.897***	1.117***	0.950***	1.144***	0.971***
	(0.135)	(0.166)	(0.0597)	(0.0804)	(0.0697)	(0.0812)	(0.0719)	(0.0813)
Age 25 - 34	0.214***	0.329***	0.491***	0.462***	0.513***	0.484***	0.530***	0.500***
	(0.0400)	(0.0257)	(0.0192)	(0.0173)	(0.0185)	(0.0167)	(0.0192)	(0.0168)
Log Wage and salary income	-0.967***	-0.602***	0.175**	0.639***	0.207***	0.727***	0.225***	0.751***
	(0.151)	(0.180)	(0.0686)	(0.122)	(0.0580)	(0.114)	(0.0591)	(0.114)
Employment rate	3.571***	0.0580	-0.0308	-1.944***	-0.204	-2.303***	-0.260	-2.372***
	(0.581)	(0.579)	(0.345)	(0.465)	(0.327)	(0.433)	(0.328)	(0.430)
Offense FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	no	no	yes	yes	yes	yes	yes	yes
Year*County FE	no	no	no	no	yes	yes	yes	yes
State*Offense FE	no	no	no	no	no	no	yes	yes
Test of equality between men and women (p-value):								
Log Earnings		0.000		0.000		0.000		0.000
Log Prob. of arrest		0.000		0.000		0.000		0.877
Observations	150,782	90,394	150,782	90,394	150,782	90,394	150,782	90,394
R-squared	0.483	0.366	0.656	0.665	0.688	0.709	0.720	0.756

Notes: The dependent variable is the log of crime rate. We use four different specifications and we estimate them separately for males and for females. Estimation includes fixed effects where noted with "yes". Excluded categories are: *white* for the race, people in the *age between 35 and 44* for the age group and *year 1995* for the year dummies. We test whether males and females have coefficients on the variables of interest that are significantly different from zero. Standard errors clustered at the county level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Log Crime rates				
	Male	Female	Male	Female	Male	Female	Male	Female
Log Illegal Earnings	0.0193***	-0.00799	0.0752***	0.0599***	0.0732***	0.0592***	0.0646***	0.0501***
	(0.00733)	(0.00728)	(0.00341)	(0.00335)	(0.00344)	(0.00344)	(0.00259)	(0.00257)
Log Prob. of arrest	-0.122***	-0.0984***	-0.122***	-0.134***	-0.102***	-0.115***	-0.0869***	-0.100***
	(0.0242)	(0.0275)	(0.0118)	(0.0115)	(0.0112)	(0.0115)	(0.0105)	(0.00962)
Black	1.379***	0.680***	1.364***	0.984***	1.334***	0.949***	1.311***	0.922***
	(0.0686)	(0.0480)	(0.0445)	(0.0365)	(0.0443)	(0.0355)	(0.0437)	(0.0346)
Age 15 - 24	0.188	-0.435**	1.018***	0.840***	1.059***	0.896***	1.089***	0.933***
	(0.140)	(0.180)	(0.0578)	(0.0788)	(0.0674)	(0.0791)	(0.0700)	(0.0802)
Age 25 - 34	0.213***	0.325***	0.489***	0.464***	0.511***	0.486***	0.530***	0.503***
	(0.0408)	(0.0273)	(0.0179)	(0.0171)	(0.0171)	(0.0164)	(0.0178)	(0.0167)
Log Wage and salary income	-0.980***	-0.719***	0.196***	0.644***	0.234***	0.734***	0.256***	0.767***
	(0.156)	(0.193)	(0.0657)	(0.120)	(0.0541)	(0.111)	(0.0557)	(0.111)
Employment rate	3.543***	0.382	-0.0394	-1.944***	-0.169	-2.314***	-0.239	-2.407***
	(0.587)	(0.605)	(0.339)	(0.460)	(0.316)	(0.422)	(0.319)	(0.421)
Offense FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	no	no	yes	yes	yes	yes	yes	yes
Year*County FE	no	no	no	no	yes	yes	yes	yes
State*Offense FE	no	no	no	no	no	no	yes	yes
Test of equality between men and women (p-value):								
Log Earnings		0.000		0.000		0.000		0.000
Log Prob. of arrest		0.000		0.000		0.000		0.000
Observations	150,782	90,394	150,782	90,394	150,782	90,394	150,782	90,394
R-squared	0.483	0.402	0.706	0.711	0.747	0.752	0.781	0.789

Notes: The dependent variable is the log of crime rate. We use four different specifications and we estimate them separately for males and for females. Estimation includes fixed effects where noted with "yes". Excluded categories are: *white* for the race, people in the *age between 35 and 44* for the age group and *year 1995* for the year dummies. We test whether males and females have coefficients on the variables of interest that are significantly different from zero. Standard errors clustered at the county level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Robustness: Other Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
			Log Crime rates			
	Male	Female	Male	Female	Male	Female
Log Illegal Earnings	0.261***	0.153***	0.424***	0.351***	0.385***	0.327***
	(0.0246)	(0.0325)	(0.0360)	(0.0515)	(0.0220)	(0.0306)
Log Probability of arrest	-0.133***	-0.131***	-0.178***	-0.122**	-0.0952***	-0.0367
	(0.0380)	(0.0438)	(0.0510)	(0.0526)	(0.0334)	(0.0369)
Controls	yes	yes	yes	yes	yes	yes
Offense FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes
Year*County FE	yes	yes	yes	yes	yes	yes
State*Offense FE	yes	yes	yes	yes	yes	yes
Observations	150,782	90,394	75,330	45,121	177.644	114,281
Robustness Check	Corrected for		Population		With Shoplifting	
	Nickell Bias		Larger than the Median			

Notes: The dependent variable is the log of crime rate. We use three different specifications and we estimate them separately for males and for females. In columns 1 and 2 we present estimates corrected for Nickell (1981) bias. In columns 3 and 4 we take a sub-sample of cohorts where the population is larger than the median. Finally, in columns 5 and 6 we present results for the whole sample, including shoplifting. Control variables include race (*black*), age groups (*age 15-24* and *age 25-34*), *log wage and salary income*, and *employment rate*. Estimation includes fixed effects where noted with "yes". Standard errors clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.