Uncovering the Gender Participation Gap in Crime

Nadia Campaniello* Evelina Gavrilova†

January 11, 2017

*Corresponding author, University of Essex and IZA, ncampa@essex.ac.uk, Wivenhoe Park, CO4 3SQ Colchester, Uk

†NHH Norwegian School of Economics, evelina.gavrilova@nhh.no, Norwegian School of Economics, Helleveien 30, 5045 Bergen, NO

Abstract

Research on the gender variation in the crime market, a peculiar labor market for illegal activities, is limited, although the issue is relevant *per se* and for its policy implications. We document a gender gap in criminal activities, based on property and white collar crimes, using data from the U.S. National Incident Based Reporting System. We show that there is a gender participation gap where around 30 percent of the crimes are committed by females. In order to explain, at least in part, the gender participation gap we investigate whether there are differences in incentives to be involved in criminal activities and in responsiveness to these incentives across gender. In particular we focus on criminal earnings and probability of arrest. We show that on average females earn 18 percent less than males while they face the same likelihood of arrest. We find that females are more responsive to changes in the expected probability of arrest, while males respond more to changes in the expected illegal earnings. The fact that females behave differently than males has implications for the heterogeneity in response to crime control policies. In addition, using a Blinder-Oaxaca type decomposition technique, we find that differences in incentives explain about 12 percent of the gender crime gap, while differences in responsiveness explain about 55 percent of the gap.

Keywords: Gender Gap in Crime, Crime Incentives, Synthetic panel
JEL Classification: J16, K42

1 Introduction

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. Indeed, the number of women committing crime worldwide is much lower than that of men, but in recent time this gender gap is shrinking. In Figure 1 we show that the trend in the percentage of women incarcerated has increased over the period 1930-2009 in the United States. The percentage of women incarcerated in 1930 was just 4.5%, increased steadily over the last 80 years, and reached 12% in 2009 [1]. The underlying participation in crime is even higher and increasing as well, as we show in Figure 2[2]. The increasing trend on both figures mimics the decrease in the gender gap in the labor market. While there is an extensive literature on the gender gap in the legal labor market, there is no economic research that looks at the gender gap in crime. Since the number of females committing crimes is on the rise, it is essential to understand how they differ from males in their criminal decision making in terms

[1]The only exception is the post-Second World War drop that followed a peak in 1945 driven by the under representation of males due to wartime service.

[2]To reconcile the participation rate of females in crime to incarceration statistics see Table A.1.

of opportunities, deterrents and violence. In this article we focus on property crimes (including white collar crimes, a quite neglected and understudied typology of crimes in the literature) because they are found to be the most responsive to incentives that, in turn, determine the expected costs and benefits of committing crimes. It stands to reason that the gender gap in the decision to engage in illegal activities might be driven on one side by biological factors and sociocultural factors like the role of women in the household, among others. On the other side, it might be driven by factors captured by the model of Becker (1968), such as opportunities in the legal labor market, economic returns to crime, deterrence and incapacitation. In our paper we focus on the latter factors in the form of incentives (and responsiveness to incentives) that alter the costs of engaging in criminal activities, thus potentially shaping the gender gap in crime.

Using the U.S. *National Incident Based Reporting System* (NIBRS) matched with population statistics for the period 1995 - 2011, we find that females earn 13 percent less than males and face a higher likelihood of being arrested than men (+4.6 percent), we also find that shoplifting is a special crime with a high participation of females. When we remove it, we find that the earnings gap increases to 18 percent but the arrest gap disappears. In order to determine how illegal earnings and probability of arrest influence the participation decision, we develop a novel empirical strategy. We use a two-stage model, where in the first stage we predict earnings and arrest with their past values, in order to identify the expectations of criminals. Then we use these predicted values to see how female and male crime rates respond to the expectations of earnings and arrest. We find that males have a higher elasticity of illegal earnings of 37 percent compared to 25 percent for females. On the other hand, females are more responsive to changes in the arrest probability with an elasticity of -15 percent compared to males at -10 percent. By exploiting a partial Blinder-Oaxaca decomposition we find that if females were more "manly" with respect to incentives and responsiveness to them this would reduce the participation gap by 67 percent.

Crucially, our elasticities with respect to the probability of arrest could be compared to previous literature. In a recent review of research, 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. Given that male and female elasticities differ significantly, we show that there would be additional heterogeneity in the response of criminals to policies that increase the likelihood of arrest. Likely, females would decrease their participation more than males.

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 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. If the policy maker wants to discourage female crime, then policies that alter the likelihood of detection and arrest are more effective. Such policies include increases in the police manpower and deployment strategies.

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 difference in the participation of males and females looking at incentives. Since then there has been scant response to this apparent gap and we are the first to fully investigate this research question. Nevertheless there is a number of recent (with the exception of Bartel (1979)) papers by economists that look at female criminals. 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.

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 $4.4 - 4.9\%$. In our paper we control for employment opportunities and wage that represent the opportunity costs of being involved in criminal activities.

Cano-Urbina *et al.* (2016) 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 in labor market opportunities.
Gavrilova (2013) finds that females are likely discriminated against in the market for criminal partnerships, which might be one of the drivers for their lower participation.
Finally, a recent paper by Miller & Segal (2014) studies the relationship between females and crime looking at victimization. The authors find that the gender composition within the police force influences female reporting and victimization rate for violent crimes.

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[3]. In line with the zeitgeist, Simon (1976) discusses the trends in female criminal behavior. Using UCR data, she notes that female crime rates have increased two times from 1932 to 1972, measured by share of females arrested.

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 (with the notable exception of Draca *et al.* (2015)), while we have more precise information on the value of the property stolen.

Finally, we provide evidence on white collar crimes. While most economic studies concentrate on property or violent crimes, as reported in the FBI annual reports, we exploit the detail provided by NIBRS and investigate white collar offenses.

2 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

[3]See for e.g. Steffensmeier & Allan (1996) for a recent summary of trends in the gender gap in crime, as seen from the perspective of sociology.

typical observation is a coded report about a criminal incident. It contains the number of perpetrators, their demographic characteristics and crime codes, a victim report on how much was stolen and an arrest report, if an arrest has been effectuated. Criminal earnings are recorded no matter whether there was an arrest.

As already mentioned, we limit our analysis to property crimes (including white collar crimes). We do this for three reasons. First, property crimes respond more to incentives than other typologies of crimes (for example violent crimes are more likely to be driven by psychological problems). Second, property crimes are more common than other crimes. The violent crime rate over the period 2011 is 387.1 crimes per 100,000 inhabitants, while the property crime rate is 2,905.4 crimes. Third, in property crimes we can observe a measure of illegal earnings.

We use the following Uniform Crime Reporting (UCR) offense codes: 231 Pocket-picking, 232 Purse-snatching, 233 Shop lifting, 234 Theft from Building, 235 Theft from Coin-Operated-Machine, 236 Theft from or of Motor Vehicle, 237 Parts, 238 All other larceny, 240 Motor Vehicle Theft, 220 Burglary, 120 Robbery, 280 Stolen Property Offenses, 261 Swindle, 262 Credit Card ATM Fraud, 263 Impersonation, 264 Welfare Fraud, 265 Wire Fraud, 250 Counterfeiting/Forgery, 270 Embezzlement, 210 Extortion/Blackmail, 510 Bribery[4].

Once we select these crime incidents we have 24 million observations on criminals over the period 1995 to 2011. A given incident can consist of several perpetrators, and we assume that if one of them was not observed well, then also the other ones are observed with measurement error. Therefore, to alleviate these measurement issues we drop in total 32 percent of the data. [5] We then select individuals between 15-44 years of age, of black or white race. This selection is guided by the availability of control variables, as we want to approximate the opportunity cost of crime in the best way we can. Therefore, we keep 40 percent of the original sample.

Ideally, to understand the participation decision we would estimate a discrete choice model of crime. However, we have information only on crimes committed, not on the other available choices. Therefore, we define "pseudo-individuals" and construct a synthetic panel (see

[4]We sum over 261 Swindle, 262 Credit Card ATM Fraud, 263 Impersonation, 264 Welfare Fraud, 265 Wire Fraud and we call them "Fraud" and we sum over 210 Extortion/Blackmail and 510 Bribery and we call them "Bribery and Extorsion"

[5]See Table A.2 for a tabulation of the missing gender variable by crime.

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. Cohorts share similar characteristics such that they would face similar incentives to commit crimes, having similar expectations over criminal proceeds and perceived likelihood of arrest. They are defined based on: interval of age (15-24, 25-34 and 35-44[6]), race (black and white), gender (male, female), and county (on average there are 2590 counties per year). The panel spans over a period of 16 years.

The resulting unbalanced panel is treated as pseudo-individuals that can be tracked over time. For some cohorts, mainly those with few observations, we sometimes get estimated probabilities of arrest that are 0 or 1. Since these are the product of cohorts of small size rather than their true expected values, we aggregate them sequentially over age group and then race, year, county, and finally typology of crime until we get values that are away from 0 or 1. As for illegal earnings in the main regressions we discard those observations where the expected illegal earnings are equal to 0, while in the robustness checks we compute them using the same strategy that we adopted for probability of arrest.

Table 1 shows the summary statistics. Across cohorts there are two times less females than males. They have both a smaller likelihood of arrest and less earnings than males.

Turning first to the participation decision, in Figure 2 we plot the ratio of the number of females to males. We show that this ratio is always below 0.5, meaning that across the years there are at least 2 men for every woman criminal. We observe that the participation of females is increasing, in line with the incarceration numbers shown in Figure 1.
In Figure 2 we present the ratio of illegal earnings of females to males. The ratio is always below one corresponding to the fact that over the sample years female earn less than males on average. This ratio is not stable across the years, which is all the more reason to concentrate on within year variation later in the empirical specification. Behind this average there might be other significant heterogeneity hidden, so we explore the density of the obtained logged illegal earnings in the top panels in Figure 3. We show that the illegal earnings of females are bimodal. In the right panel we drop shoplifting and show that the illegal earnings of males

[6]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.

and females become more similar.

When we look at the arrest ratio in Figure 2 we observe that females are more likely to be arrested than males. When we look at the density of the log of the arrest rate Figure 3, we observe that again shoplifting is biasing the distribution of observed female criminals when we compare the two figure in the bottom panels. It seems that shoplifting is observed often when there is an arrest, which translates into many zeros in the log of probability of arrest. Once we drop shoplifting, the distributions of males and females are more similar.

In Figure 4 we show how participation varies with respect to illegal earnings and arrest for different crime types. The size of the circle is proportional to the fraction of females in each crime and it is noticeable that females participate most in shoplifting. In addition, when we show the linear fit between gender participation gap and illegal earnings (in log, top panel), and gender participation gap and arrest (in log, bottom panel) we observe that shoplifting significantly biases the relationship (in the left part of the panel we plot all property crimes, in the right one we exclude shoplifting). Since shoplifting represents a clear outlier, in our main results we get rid of it and we include it in the robustness checks.

Finally, in Table 1 we present the control variables. 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 employment rates (by age, gender, race, decade, and state) are taken from the CENSUS Integrated Public Use Microdata Series (IPUMS-USA).

3 Model of Crime

3.1 Theory

In this section we present a model of crime participation, based on Becker (1968). An individual compares the expected utility of committing a crime with the expected utility of not committing a crime. The expected benefits of committing a crime are the illegal earnings while the expected costs are the probability of arrest and the sanction length, and the opportunity costs of being engaged in the legal labor market (wage, employment rate).

In a generalized version, an individual decides to be involved in a criminal activity if a function of costs and benefits is larger than an individual idiosyncratic error Υ^g, which measures any unobserved determinants of crime (sociocultural reasons, family reasons, biological and genetic factors, etc.) for gender $g = m, f$:

$$f^g(\widehat{Y}^g, \widehat{P}^g, \widehat{WAGES}^g, \widehat{EMPLOYMENT}^g, \widehat{SANCTION}^g) > \Upsilon^g \quad , \tag{1}$$

where $\widehat{hats}$ indicate expectations. So, for example, $\widehat{Y}^g$ are expected economic returns to crime (criminal earnings) and $\widehat{P}^g$ is the expected probability of arrest.

In order to derive an estimable equation we do the following steps. First we aggregate the equation across individuals, deriving crime rates. Second, since the function $f()$ is unknown, we log-linearize $f()$ with respect to all the incentives variables (small letters for logs):

$$crimerate^g = \beta_1^g \widehat{y}^g + \beta_2^g \widehat{p}^g + \beta_3^g \widehat{wages}^g + \beta_4^g \widehat{employment}^g + \beta_5^g \widehat{sanction}^g + \varepsilon^g \tag{2}$$

The equation shows that the gender crime gap could be due to differences in incentives as well as differences in the way criminals of different genders respond to such incentives.

3.2 Estimation

In this section we present our empirical methodology. First, we map out 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 Figure 2 we documented significant differences between males and females in terms of illegal earnings, arrest and participation. To show how these differences vary with other factors that are also correlated with gender we estimate the following specification:

$$incentive_{igt} = \beta 1(g = f)_{igt} + X'_{igt}\gamma + \varepsilon_{igt} \tag{3}$$

where $incentive_{it}$ is either the log arrest rate or the log transformed value of property stolen of criminals in cohort i and year t and $1(g = f)$ is an indicator function equal to one when the gender is female. X is a vector containing personal-cohort traits like race, age, average wage and unemployment rate. A $\beta = 0$ would imply that for the baseline that there is no gender gap in any of these measures.

Including offense fixed effects allows us to determine how much of the unconditional gap is due to differences between offenses. 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, in some specifications the error term includes county-year fixed effects. Furthermore in the last specifications the error term 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.

3.2.2 Responsiveness to Incentives

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

$$crimerate_{it}^{g} = \beta_1^g \widehat{y}_{it}^g + \beta_2^g \widehat{p}_{it}^g + X_{it}^{g\prime} \delta^g + \varepsilon_{it}^g, \text{ for } g = m, f, \tag{4}$$

where *crimerate* is the log crime rate defined as the number of crimes committed in a given year by people within a cohort, defined by age group, race, and county, divided by the general population in the same cohort. $\widehat{y}$ is the expected log of illegal earnings, $\widehat{p}$ is the expected log probability of being apprehended. X is a vector containing personal traits (wage and salary income, employment rate, race, age groups). As for Equation 3 in some specifications the error term includes county-year fixed effects, in order to control for county specific heterogeneity such as police presence in any given year. In the last specifications the error term includes 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 this equation separately for males and females.

Using the contemporaneous values $\widehat{y}_{i,t}$ and $\widehat{p}_{i,t}$ we face two potential issues 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.

One solution would be to use the lagged values of the incentives, which implicitly assumes that the proper expectations of criminals are adaptive, being based on what happened in the previous year. Since it might be the case that $z_{i,t-1} = (y_{i,t-1}, p_{i,t-1})$ are not the true criminals' expectations, we face potential measurement error. Rather than relying on such an assumption, we model the expectations for both illegal earnings and probability of arrest and test whether they are adaptive (α=1):

$$\widehat{z_{it}^{g}} = \alpha^{g} z_{i,t-1}^{g} + X_{it}' \rho^{g} + \xi_{it}^{g}, \text{ for } g = m, f. \tag{5}$$

In modeling the expectations we use a two step procedure, where in the first stage we obtain $\widehat{p}$ and $\widehat{y}$ and in the second step these measures are plugged into equation 4. In other words, these equations are similar to a first stage in a 2SLS setup where y_{t-1} and p_{t-1} are used as instruments for y_t and p_t.

3.3 Blinder-Oaxaca decomposition

In order to gauge the importance of the incentives and their elasticities in determining the gender crime gap we use a partial Blinder-Oaxaca decomposition (limited to p and y). The decomposition measures the fraction of the gender crime gap that arises because females and males, on average: i) face different incentives and ii) respond differently to incentives.

The counterfactual equation for women where we replaced their incentives (their "endowment') with those from male equation is:

$$\widehat{crimerate}_{fit}^{CF,Xs} = \widehat{crimerate}_{fit} + (z_{ict}^{m} - z_{it}^{f})\widehat{\beta}_{z}^{f} \tag{6}$$

The counterfactual equation for women where we replace their coefficients on incentives with those from the male equation is:

$$\widehat{crimerate}_{fit}^{CF,\beta s} = \widehat{crimerate}_{fit} + (\widehat{\beta}_z^m - \widehat{\beta}_z^f)(z_{it}^f - \underline{z}^f). \tag{7}$$

Subtracting the minimum value $\underline{z}^f$ normalizes the intercept and rotates the counterfactual around its minimum value (rather than around $z = 0$).

It can be shown the fraction of the crime gap that can be explained by differences in incentives and elasticities is

$$\frac{\widehat{crimerate}_{fit}^{CF,\beta s} - \widehat{crimerate}_{fit}}{\widehat{crimerate}_{mit} - \widehat{crimerate}_{fit}} + \frac{\widehat{crimerate}_{fit}^{CF,Xs} - \widehat{crimerate}_{fit}}{\widehat{crimerate}_{mit} - \widehat{crimerate}_{fit}} \tag{8}$$

4 Results

4.1 Differences in incentives

In Table 2 we present results for the illegal earnings gap. As we explain in Section 2, shoplifting is a special crime, therefore we exclude it from the first four columns and we include it in the last column as a robustness check. In column 1 we show that conditional on demographic characteristics and variables that affect the opportunity cost of crime, females have 18 percent less criminal earnings. In the next three columns we add progressively more fixed effects, yet the gap is stable around 18 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. In column 5 we observe that the gap is still negative and larger in magnitude. Females earn 22 percent less than males in daylight crimes. This implies that, if anything, reporting bias would drive the gender gap towards 0.

As a robustness check we include shoplifting in the sample in column 6. We find that the gap diminishes to 13 percent meaning that the earnings gap in shoplifting is in favor of females and is significant enough to attenuate estimates on the earnings gap in all crimes.

In Table 3 we present the results for the arrest gap. In the first column the arrest gap is negative, but it turns insignificant as we control for more heterogeneity in different counties and years and we control for the typology of offense they commit and the interaction between offense and State fixed effects. As before, our preferred in specification is that in column 4. The arrest gap remains insignificant also in daylight crimes, suggesting that the result is not driven by reporting bias. In the last column we show that shoplifting drives the estimates on the average gap in arrest rates, which now flips to 4.6 percentage points, such that females face a higher likelihood of arrest. In line with what we already shown the reason is that, in the grand majority of cases, shoplifting (the most committed crime by females) is reported when the criminal gets arrested thus upward biasing the coefficient on arrest for females.

Therefore, conditional on crime participation we find that females earn 13 to 18 percent less than males and face the same likelihood of being arrested (when we do not include shoplifting).

4.2 Differences in Responsiveness

Table 4 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. The lag of the (log) probability of arrest and the lag of (log) illegal earnings are good predictors of, respectively, the probability of arrest and illegal earnings. The F-statistics is, in all the specifications, well above conventional levels. Since the αs in both the equations are lower than 1, it follows that expectations are not adaptive meaning that we should take the instrumental variable as our preferred specification. In line with the first stage results, the reduced form cannot be interpret as the true elasticities but 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 are associated with a decrease in the crime rates.

In Table 5 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 8 to 37 percent higher participation of males in crime. Similarly, the elasticity of expected criminal earnings for females is between from 0 to 32 percent, with a final coefficient of 25 percent in our preferred specification in the last column. On average, we observe that females have always a significantly lower elasticity of illegal earnings than males.

With respect to the probability of arrest, we find that the elasticity of females is between -10 and -22 percent, while that of males is between -10 and -16 percent. In our preferred specification (columns 7 and 8, the male coefficient is -10 percent and the female one is -15 percent). Therefore, females respond significantly stronger to an increase in the probability of apprehension than males do.

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. When we use the fixed effects the employment rate is negatively associated, as expected, with the crime rates, while average wage is positively associated with crime rates (it might be explained thinking that income reflects the presence of individuals who provide good targets for criminals involved in property crimes). To sum up, when modeling the expectations we find that both males and females respond to incentives. Females are more responsive to fluctuations in the probability of arrest, while males are more responsive to the changes in illegal earnings.

In the appendix, in Table A.3, we show OLS estimates. We find that the estimates of the elasticities are smaller and sometimes fail to be significant. These results are consistent with the measurement error and the endogeneity issues outlined before.

Robustness Checks. In Table 6 we perform some robustness checks to be sure that our results do not depend on the particular specification we used. First of all, instead of discarding data with missing values for illegal earnings, we assign them the average of illegal earnings that we compute aggregating over age group and then race, year, county, and finally typology of crime until we get values that are different from 0. The differences between males and females remain the same, though the elasticities are smaller for illegal earnings and larger for arrest.

To be sure that our results are not biased by the different dimension of the counties, we weight the estimates by the population in the county. Estimates remain similar. When we add shoplifting, we find that the elasticities are very similar to those of our main specification for earnings but lower in magnitude for arrest. Finally when using the number of crimes instead of the crime rates as our dependent variable, our estimates are in line with our main results even though the elasticities on illegal earnings are larger in magnitude.

4.3 Blinder-Oaxaca decomposition

Using Equation 8 the fact that women respond more to the arrest probabilities and less to the monetary incentives explains 55 percent of the gap. Another 12 percent is due to the differences in incentives men and women face. The counterfactual CDFs of female crime if they responded to the incentives like men are shown in Figure 6, while Figure 5 shows what would happen if they faced the same incentives.

5 Conclusion

In this article we reveal that gender gaps are not only a feature of the labor market and we contribute to the economic literature on crime. We find that females are less likely to engage in crime (they represent 30 percent of the total property crimes in the United States). To identify the possible motives for the gender gap in crime participation we look at possible differences in incentives to engage in criminal activities across gender (we focus on illegal earnings and probability of arrest) and we investigate whether males and females respond differently to perceived incentives to commit crimes. We find that female criminals earn 18 percent less than male criminals while they face the same likelihood of arrest. Such difference in incentives explains 12 percent of the gender gap in crime. We find that the decision of both males and females to engage in criminal activities depends on the expected earnings and on the probability of being arrested. Men are significantly more responsive than women to the expected monetary incentives, while women show a higher disutility to get arrested (probably for some sociocultural factors like their role in the households, child-rearing, social norms, etc.). Such differences in incentives and in the way the two genders respond to incentives explains 55 percent of the male-female participation gap in crime.

This paper has focused on trying to explain the gender participation gap in crime while has not tackled the issue of the gender convergence over time that we leave for future investigation. Finally, we leave the resolution of the participation gap that we could not explain for future research.

6 Acknowledgements

We are grateful to Giovanni Mastrobuoni, Rodrigo R. Soares, Libor Dusek, Asaf Zussman, Michele Pellizzari, Olivier Marie, Mirko Draca and David Card for their useful comments. We also thank all the participants to the Essex internal seminar, Colchester, UK; the Amsterdam Centre for Law and Economics seminar, Amsterdam, the Netherlands, the First Workshop on Culture and Institutions, Galatina, Italy; the First Conference on Empirical Legal Studies in Europe (CELSE), Amsterdam, Netherlands; the 30th Annual Conference of the European Society for Population Economics, Berlin, Germany; the Stockholm Criminology Symposium, Stockholm, Sweden; the Alpine Population Conference, Dondena Centre, Bocconi University, La Thuile, Italy; the Italian Society of Law and Economics - 10th Annual Conference, Roma, Italy; the XII Brucchi Luchino Labor Economics Workshop, Napoli, Italy; the 26th annual Conference of the European Association of Labor Economists, Ljubljana, Slovenia; the 13th Journes d'Economie Publique Louis-Andrè Grard-Varet, Aix-en-Provence, France; the Sixth Annual Meeting on the Economics of Risky Behaviors (IZA), Medellin, Colombia; the Scottish Economic Society Annual Conference, Perth, Scotland.

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.*

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

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.*

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.

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?

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

Gavrilova, Evelina. 2013. A Partner in Crime: Assortative Matching and Bias in the Crime Market.

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

Miller, Amalia R, & Segal, Carmit. 2014. Do female officers improve law enforcement quality? Effects on crime reporting and domestic violence escalation. *Effects on Crime Reporting and Domestic Violence Escalation (September 16, 2014)*.

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, & Allan, Emilie. 1996. Gender and Crime: Toward a Gendered Theory of Female Offending. *Annual Review of Sociology*, **22**(1), 459–487.

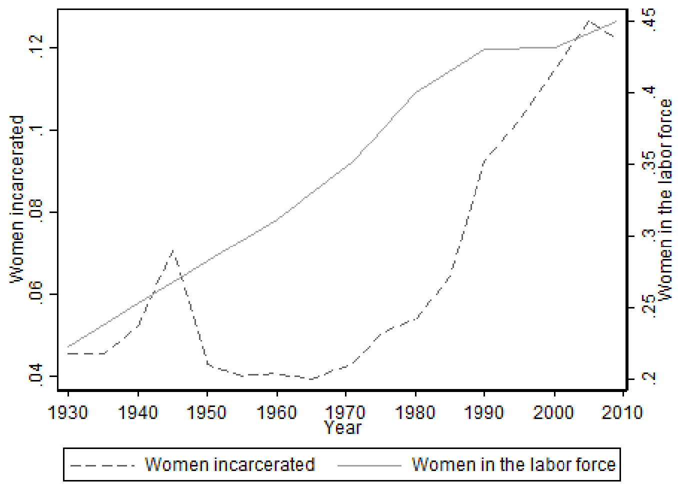


Figure 1: Fraction of Women Incarcerated and in the Labour Force in the U.S

Notes: On the left axis we plot the fraction of females incarcerated 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.

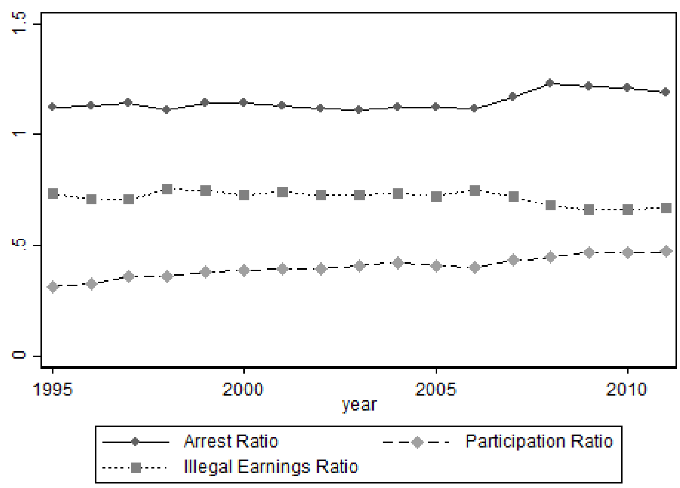
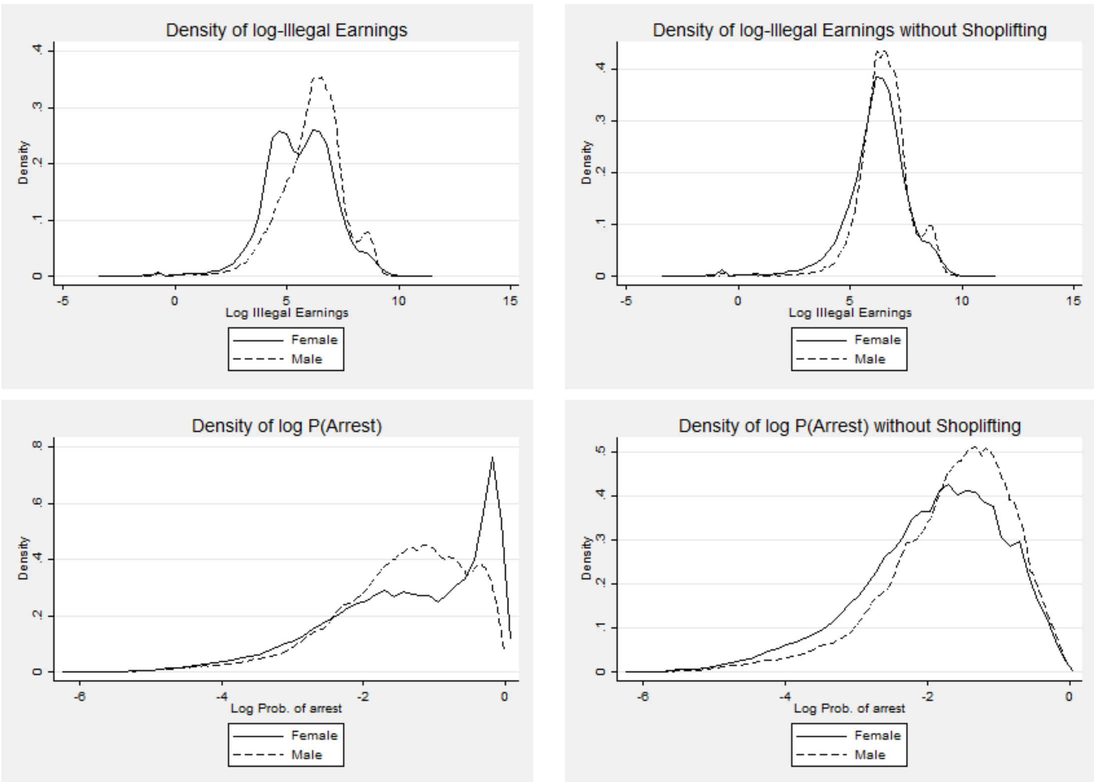


Figure 2: Ratios of Females to Males

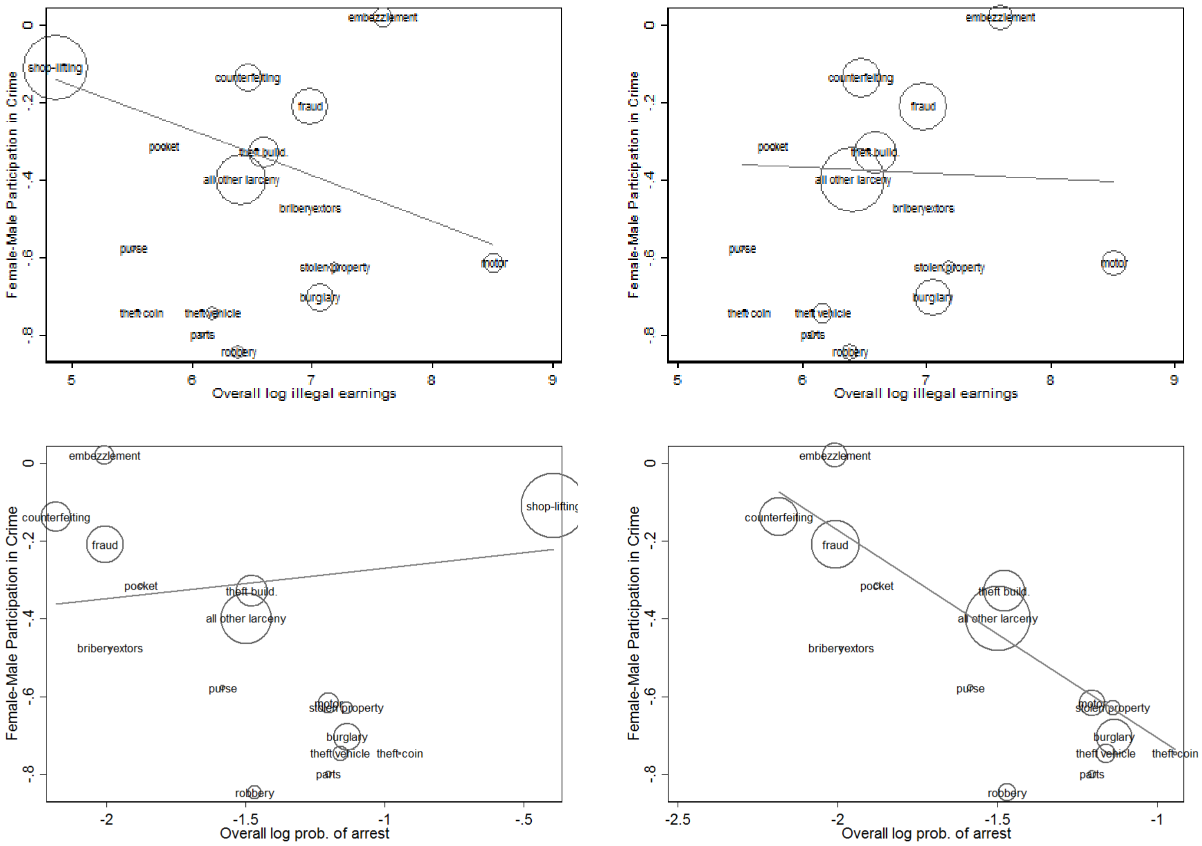
Notes: In this graph the relative participation rate, arrest rate and earnings of females are plotted with respect to time. Each data point represents the ratio of female criminals to male criminals in a given year.

Figure 3: Density of Illegal Earnings and Arrest for Males and Females



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 4: Gender Participation Gap in Crime and Incentives



Notes: Here we plot the relationship between gender crime gap and average incentives across the two gender (in log).
Y-axis: $\frac{FemaleCrimes - MaleCrimes}{FemaleCrimes + MaleCrimes}$.
X-axis: $\frac{MaleCrimes}{FemaleCrimes + MaleCrimes} * log(Ill.EarningsofMales) + \frac{FemaleCrimes}{FemaleCrimes + MaleCrimes} * log(Ill.EarningsofFemales)$.
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.

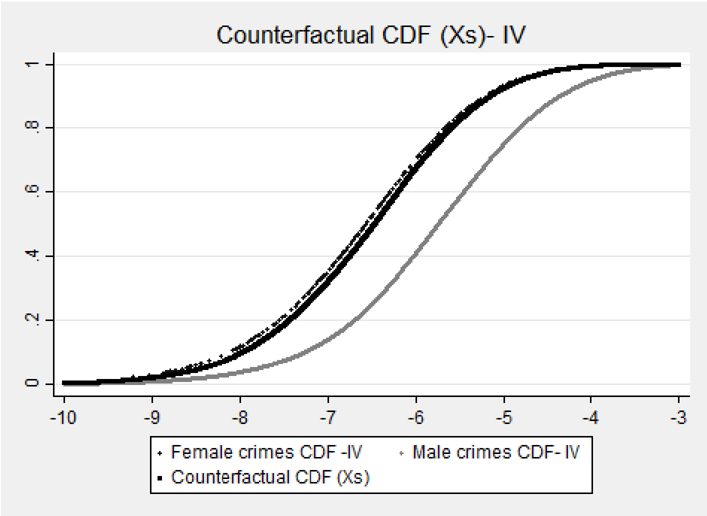


Figure 5: Crime Gaps in the counterfactual scenario - Males Xs in the Female Equation

Notes: We plot 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.

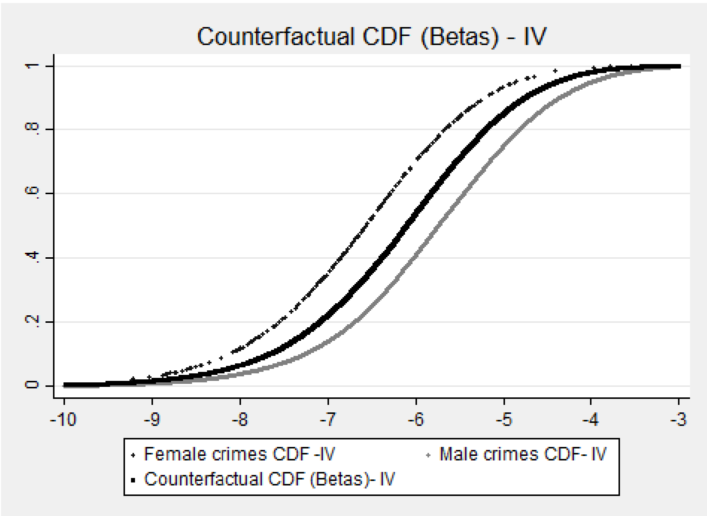


Figure 6: Crime Gaps in the counterfactual scenario - Males βs in the Female Equation

Notes: We plot 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: Summary Statistics for the Synthetic Panel

	Males					Females				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Number of crimes	130,905	31.411	76.734	1	4862	116,369	12.141	19.590	1	437
Prob. of arrest	130,905	0.266	0.189	0.002	0.985	116,369	0.236	0.174	0.003	0.971
Illegal earnings	130,905	1345.738	2127.497	0.042	48000	116,369	1277.677	2437.313	0.050	55918
Population	130,905	10441.070	17091.890	11	250451	116,369	10583.730	16670.850	3	231350
Wage and salary income	130,905	21490.790	13417.040	940	83930.83	116,369	13878.190	7202.535	1392.857	54956.44
Employment rate	130,905	0.700	0.169	0	1	116,369	0.647	0.106	0.250	1
Age 15 - 24	130,905	0.393	0.488	0	1	116,369	0.382	0.486	0	1
Age 25 - 34	130,905	0.339	0.473	0	1	116,369	0.343	0.475	0	1
Black	130,905	0.276	0.447	0	1	116,369	0.273	0.445	0	1
Year	130,905	2005.446	4.018	1996	2011	116,369	2005.448	4.002	1996	2011
Embezzlement	130,905	0.057	0.232	0	1	116,369	0.064	0.245	0	1
Fraud	130,905	0.147	0.354	0	1	116,369	0.160	0.367	0	1
Counterfeiting	130,905	0.039	0.193	0	1	116,369	0.041	0.197	0	1
Bribery and Extortion	130,905	0.000	0.015	0	1	116,369	0.000	0.013	0	1
Pocket-picking	130,905	0.005	0.068	0	1	116,369	0.005	0.070	0	1
Purse-snatching	130,905	0.005	0.068	0	1	116,369	0.004	0.065	0	1
Theft from building	130,905	0.136	0.343	0	1	116,369	0.142	0.349	0	1
Theft from coin operated machine	130,905	0.001	0.030	0	1	116,369	0.001	0.024	0	1
Theft from a motor vehicle and its parts	130,905	0.067	0.249	0	1	116,369	0.064	0.245	0	1
Parts	130,905	0.015	0.121	0	1	116,369	0.012	0.111	0	1
All other larceny	130,905	0.237	0.425	0	1	116,369	0.248	0.432	0	1
Motor vehicle theft	130,905	0.097	0.296	0	1	116,369	0.093	0.290	0	1
Burglary	130,905	0.141	0.348	0	1	116,369	0.128	0.334	0	1
Robbery	130,905	0.043	0.203	0	1	116,369	0.034	0.180	0	1
Stolen property offenses	130,905	0.012	0.107	0	1	116,369	0.005	0.068	0	1

Notes: 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.

Table 2: Differences in Illegal Earnings

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Illegal Earnings					
Female	-0.180***	-0.190***	-0.179***	-0.180***	-0.218***	-0.130***
	(0.0331)	(0.0235)	(0.0103)	(0.00964)	(0.0112)	(0.00861)
Black	0.125***	-0.0246	-0.0242	-0.0423**	-0.0697***	0.0685***
	(0.0371)	(0.0308)	(0.0222)	(0.0203)	(0.0204)	(0.0178)
Age 15 - 24	-0.0693	-0.109***	-0.188***	-0.182***	-0.195***	-0.129***
	(0.0660)	(0.0318)	(0.0307)	(0.0270)	(0.0311)	(0.0275)
Age 25 - 34	0.00384	-0.0143	-0.0158	-0.0123	-0.0200*	0.0268***
	(0.0171)	(0.0130)	(0.0103)	(0.00993)	(0.0112)	(0.00951)
Log Wage and salary income	0.416***	0.283***	0.197***	0.190***	0.172***	0.189***
	(0.0688)	(0.0395)	(0.0364)	(0.0351)	(0.0364)	(0.0359)
Employment rate	-0.968***	-0.528***	-0.392***	-0.346***	-0.255*	-0.154
	(0.257)	(0.151)	(0.126)	(0.128)	(0.135)	(0.150)
Log Population	0.218***	0.115***	0.126***	0.115***	0.112***	0.108***
	(0.0126)	(0.0174)	(0.0117)	(0.00889)	(0.00977)	(0.00807)
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	247,274	247,274	247,274	247,274	182,708	283,750
R-squared	0.087	0.204	0.526	0.557	0.495	0.692

Notes: The dependent variable in all regressions is the log of illegal earnings. 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 3: Differences in Probability of Arrest

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Probability of arrest					
Female	-0.204***	-0.268***	-0.0170	-0.0186	0.0215	0.0458***
	(0.0636)	(0.0578)	(0.0135)	(0.0133)	(0.0143)	(0.0101)
Black	-0.166***	-0.0696**	-0.00806	-0.0172	0.0274	-0.0521**
	(0.0524)	(0.0328)	(0.0263)	(0.0245)	(0.0272)	(0.0206)
Age 15 - 24	0.767***	0.588***	0.543***	0.528***	0.460***	0.426***
	(0.110)	(0.0759)	(0.0434)	(0.0420)	(0.0472)	(0.0319)
Age 25 - 34	-0.00279	-0.0260	-0.0397*	-0.0408**	-0.0440**	-0.0602***
	(0.0342)	(0.0229)	(0.0204)	(0.0201)	(0.0219)	(0.0191)
Log Wage and salary income	0.303**	0.175**	0.173***	0.172***	0.133**	0.166***
	(0.132)	(0.0787)	(0.0502)	(0.0519)	(0.0588)	(0.0444)
Employment rate	0.109	0.180	0.480***	0.425**	0.381*	0.224
	(0.434)	(0.219)	(0.186)	(0.183)	(0.213)	(0.166)
Log Population	-0.0700***	-0.0281	-0.0351**	-0.0317**	-0.0202	-0.0156
	(0.0207)	(0.0181)	(0.0138)	(0.0144)	(0.0157)	(0.0122)
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	247,274	247,274	247,274	247,274	182,708	283,750
R-squared	0.147	0.308	0.537	0.578	0.555	0.685

Notes: The dependent variable in all regressions is the log of the probability of arrest. 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: 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	0.320***	0.236***	0.152***	0.113***	0.139***	0.107***	0.102***	0.0740***
	(0.0249)	(0.0188)	(0.00782)	(0.00630)	(0.00582)	(0.00534)	(0.00504)	(0.00453)
Log Lag Prob. of arrest	-0.0329***	-0.0457***	-0.0548***	-0.0397***	-0.0565***	-0.0375***	-0.0308***	-0.0221***
	(0.0117)	(0.0124)	(0.00751)	(0.00788)	(0.00793)	(0.00793)	(0.00679)	(0.00723)
				Log Probability of arrest				
Log Lag Prob. of arrest	0.509***	0.463***	0.362***	0.305***	0.368***	0.311***	0.323***	0.268***
	(0.00901)	(0.01000)	(0.00990)	(0.00913)	(0.0105)	(0.00950)	(0.00987)	(0.00914)
Log Lag Illegal Earnings	-0.00101	-0.00108	-0.00828***	-0.00239	-0.00895***	-0.00186	-0.00347*	0.000157
	(0.00303)	(0.00238)	(0.00207)	(0.00171)	(0.00216)	(0.00172)	(0.00187)	(0.00164)
First stage F-stat	80.78	77.24	192.8	164.9	298.5	200.8	203.2	133.9
				Panel B: Reduced form				
				Log Crime rates				
Log Lag Illegal Earnings	0.0259***	-0.0001	0.0560***	0.0367***	0.0530***	0.0353***	0.0383***	0.0187***
	(0.00676)	(0.00555)	(0.00323)	(0.00278)	(0.00324)	(0.00280)	(0.00248)	(0.00215)
Log Lag Probability of arrest	-0.0738***	-0.0495***	-0.0770***	-0.0785***	-0.0703***	-0.0697***	-0.0440***	-0.0467***
	(0.0155)	(0.0168)	(0.00686)	(0.00716)	(0.00706)	(0.00737)	(0.00629)	(0.00668)
Controls	yes	yes	yes	yes	yes	yes	yes	yes
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		0.381	
Observations	130,905	116,369	130,905	116,369	130,905	116,369	130,905	116,369

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*. 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". 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 5: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Crime rates							
	Male	Female	Male	Female	Male	Female	Male	Female
Log Earnings	0.0804***	-0.001	0.359***	0.321***	0.371***	0.328***	0.374***	0.253***
	(0.0218)	(0.0235)	(0.0273)	(0.0277)	(0.0236)	(0.0265)	(0.0272)	(0.0306)
Log Prob. of arrest	-0.140***	-0.107***	-0.158***	-0.216***	-0.134***	-0.185***	-0.101***	-0.153***
	(0.0302)	(0.0363)	(0.0188)	(0.0227)	(0.0191)	(0.0233)	(0.0191)	(0.0243)
Black	1.411***	0.875***	1.287***	1.059***	1.218***	1.023***	1.201***	0.985***
	(0.0678)	(0.0422)	(0.0492)	(0.0376)	(0.0518)	(0.0367)	(0.0516)	(0.0362)
Age 15 - 24	0.191	-0.230	1.028***	1.078***	1.019***	1.131***	1.002***	1.101***
	(0.138)	(0.157)	(0.0569)	(0.0777)	(0.0703)	(0.0817)	(0.0732)	(0.0805)
Age 25 - 34	0.199***	0.321***	0.492***	0.483***	0.509***	0.498***	0.514***	0.501***
	(0.0389)	(0.0231)	(0.0183)	(0.0148)	(0.0183)	(0.0153)	(0.0187)	(0.0153)
Log Wage and salary income	-1.099***	-0.833***	0.154**	0.722***	0.197***	0.812***	0.190***	0.783***
	(0.157)	(0.176)	(0.0697)	(0.108)	(0.0646)	(0.107)	(0.0654)	(0.106)
Employment rate	3.642***	1.372***	-0.499	-1.901***	-0.816**	-2.223***	-0.858**	-2.152***
	(0.579)	(0.512)	(0.347)	(0.412)	(0.340)	(0.399)	(0.342)	(0.394)
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	yes	yes	yes	yes	yes	yes	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		0.026	
Observations	130,905	116,369	130,905	116,369	130,905	116,369	130,905	116,369
R-squared	0.485	0.414	0.614	0.580	0.665	0.635	0.697	0.704

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 6: Robustness: Other Specifications

	(1) Males	(2) Females	(3) Males	(4) Females	(5) Males	(6) Females	(7) Males	(8) Females
				Log Crime rates				
WITH MISSING DATA								
Log Lag Illegal Earnings	0.0751***	-0.0198	0.270***	0.235***	0.254***	0.232***	0.249***	0.175***
	(0.0245)	(0.0288)	(0.0294)	(0.0305)	(0.0281)	(0.0290)	(0.0264)	(0.0277)
Log Lag Probability of arrest	-0.172***	-0.121***	-0.244***	-0.309***	-0.216***	-0.268***	-0.157***	-0.206***
	(0.0267)	(0.0331)	(0.0159)	(0.0191)	(0.0163)	(0.0199)	(0.0158)	(0.0196)
WIGHTED BY POPULATION								
Log Lag Illegal Earnings	0.153***	0.194***	0.376***	0.424***	0.326***	0.370***	0.332***	0.336***
	(0.0566)	(0.0501)	(0.0514)	(0.0470)	(0.0439)	(0.0387)	(0.0378)	(0.0417)
Log Lag Probability of arrest	-0.279***	-0.291***	-0.189***	-0.204***	-0.177***	-0.182***	-0.134***	-0.130***
	(0.0623)	(0.0835)	(0.0311)	(0.0407)	(0.0305)	(0.0391)	(0.0325)	(0.0372)
WITH SHOPLIFTING								
Log Lag Illegal Earnings	0.0902***	0.0189	0.362***	0.349***	0.370***	0.357***	0.381***	0.299***
	(0.0216)	(0.0232)	(0.0244)	(0.0261)	(0.0202)	(0.0244)	(0.0238)	(0.0281)
Log Lag Probability of arrest	-0.119***	-0.0777**	-0.116***	-0.158***	-0.102***	-0.142***	-0.0693***	-0.108***
	(0.0301)	(0.0358)	(0.0174)	(0.0194)	(0.0167)	(0.0202)	(0.0179)	(0.0218)
				Log Number of Crimes				
WITH NUMBER OF CRIMES AS DEP. VARIABLE								
Log Lag Illegal Earnings	0.278***	0.297***	0.405***	0.400***	0.427***	0.419***	0.454***	0.390***
	(0.0328)	(0.0348)	(0.0291)	(0.0295)	(0.0238)	(0.0269)	(0.0288)	(0.0333)
Log Lag Probability of arrest	-0.246***	-0.284***	-0.152***	-0.196***	-0.130***	-0.171***	-0.0983***	-0.139***
	(0.0251)	(0.0268)	(0.0188)	(0.0224)	(0.0192)	(0.0233)	(0.0188)	(0.0242)
Controls	yes	yes	yes	yes	yes	yes	yes	yes
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

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. Control variables include race (*black*), age groups (*age 15-24* and *age 25-34*), *log wage and salary income*, and *employment rate*. 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". Standard errors clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

A Appendix Tables

The participation rate of females is difficult to reconcile to incarceration statistics. It is well known that for every ten incarcerated males there is one female. We perform a meta analysis on the way of females into the justice system in Table A.1. Each column "Females" refers to the fraction of females that first committed the crime, then got arrested, served as defendant and got incarcerated. We compare this fraction across each step of the judicial process. We consistently observe that females fall out of the criminal justice system, as their relative fractions decreases. Therefore, we find that the incarceration statistics are a biased representation of the participation of females in crimes. With that in mind, our focus is going to be on the decision to engage in crime and on how such decision depends on the probability of arrest.

Table A.1: Reconciling Crime Reports and Arrest Figures

	Crime	Females	Arrested	Females	Defendants	Females	Incarcerated	Females
Larceny	55 %	38 %	61 %	44 %	36 %	31 %	20 %	17 %
Burglary	12 %	15 %	10 %	12 %	37 %	11 %	60 %	5 %
Motor Vehicle	3 %	20 %	3 %	17 %	11 %	16 %	7 %	6 %
Others	27 %	29 %	25 %	28 %	15 %	17 %	13 %	1 %

Notes: The second column shows what percentage of property crimes is classified according to the crime categories in the first column. The third column shows what percentage of these crimes were committed by females. The fourth column shows from all of the arrested how much were arrested for each crime category, with respectively how much of these arrests were females. The sixth and seventh column show how much of all incarcerated were put behind bars for each of the crime offenses and what percentage of that were females. All statistics pertain to the year 2010, except for defendants - 2009. Source: NIBRS, 2010, Department of Justice, 2012

Table A.2: Unknown Offenders By Crime

UCR	Crime	Mean	SD
231	Pocket-picking	0.11	0.02
232	Purse-snatching	0.08	0.02
233	Shoplifting	0.01	0.00
234	Theft From Building	0.12	0.03
235	Theft From Coin-Operated Machine	0.20	0.06
236	Theft from Motor Vehicle	0.17	0.04
237	Parts	0.21	0.06
238	All Other Larceny	0.13	0.03
240	Motor Vehicle Theft	0.14	0.04
220	Burglary	0.15	0.03
120	Robbery	0.04	0.01
280	Stolen Property Offenses	0.04	0.01
261	Swindle	0.07	0.02
262	Credit Card ATM Fraud	0.11	0.03
263	Impersonation	0.09	0.04
264	Welfare Fraud	0.03	0.02
265	Wire Fraud	0.16	0.04
250	Counterfeiting Forgery	0.07	0.02
270	Embezzlement	0.02	0.01
210	Extortion Blackmail	0.04	0.02
510	Bribery	0.03	0.02
	Average	0.10	

Notes: This table presents the fraction of offenders for whom the gender is not known. The Mean and Standard Deviation are defined over the time series of the sample period 1995 - 2011.

Table A.3: OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Crime rates							
	Male	Female	Male	Female	Male	Female	Male	Female
Log Earnings	0.0498***	0.0287***	0.0851***	0.0716***	0.0830***	0.0705***	0.0686***	0.0550***
	(0.00694)	(0.00566)	(0.00350)	(0.00289)	(0.00335)	(0.00289)	(0.00260)	(0.00219)
Log Prob. of arrest	-0.0616***	-0.0188	-0.0618***	-0.0401***	-0.0468***	-0.0230***	-0.0234***	-0.00190
	(0.0156)	(0.0166)	(0.00767)	(0.00742)	(0.00743)	(0.00724)	(0.00708)	(0.00661)
Black	1.406***	0.886***	1.207***	0.990***	1.137***	0.950***	1.110***	0.925***
	(0.0672)	(0.0419)	(0.0482)	(0.0344)	(0.0502)	(0.0333)	(0.0498)	(0.0325)
Age 15 - 24	0.151	-0.281*	0.912***	0.985***	0.906***	1.034***	0.895***	1.021***
	(0.141)	(0.163)	(0.0533)	(0.0730)	(0.0673)	(0.0758)	(0.0704)	(0.0762)
Age 25 - 34	0.200***	0.312***	0.484***	0.477***	0.501***	0.494***	0.509***	0.499***
	(0.0393)	(0.0238)	(0.0170)	(0.0140)	(0.0171)	(0.0143)	(0.0178)	(0.0146)
Log Wage and salary income	-1.079***	-0.875***	0.161**	0.709***	0.201***	0.786***	0.201***	0.762***
	(0.160)	(0.183)	(0.0679)	(0.103)	(0.0629)	(0.100)	(0.0646)	(0.102)
Employment rate	3.501***	1.402***	-0.556	-1.894***	-0.844**	-2.181***	-0.899***	-2.134***
	(0.581)	(0.527)	(0.348)	(0.396)	(0.337)	(0.378)	(0.340)	(0.381)
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	yes	yes	yes	yes	yes	yes	yes	yes
Test of equality between males and females (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	130,905	116,369	130,905	116,369	130,905	116,369	130,905	116,369

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.