

# **ECON 5140 Applied Econometrics**

## **Semester Paper**

**Satya Harish Reddy Pulipelli**

### **INTRODUCTION**

Crime victimization represents a significant social and economic burden on individuals and communities across the United States. However, official crime statistics often fail to capture the full extent of criminal activity, as many incidents go unreported to law enforcement. This reporting gap not only distorts our understanding of crime patterns but may also exacerbate existing social inequalities if certain demographic groups face greater barriers to accessing justice through formal reporting channels.

This study examines a critical question: "Are women and minorities in low-wage jobs less likely to report crimes, and does this effect vary between urban and rural areas?" Understanding differential reporting patterns across socioeconomic and demographic dimensions is vital for both policy development and theoretical advancement. If systematic disparities exist in crime reporting behavior, vulnerable populations may experience compounded victimization—first by the crime itself and second by their inability to access protective services.

Building on previous research, I investigated two key hypotheses. First, I propose that low-wage workers in urban areas are less likely to report crimes than their rural counterparts, potentially due to differences in police-community relations and resource constraints. Second, I hypothesize that women and minorities in low-wage jobs report crimes at lower rates compared to non-minorities

and men, possibly stemming from financial constraints, fear of retaliation, and historical patterns of institutional distrust.

Using data from the National Crime Victimization Survey (NCVS), this analysis employs linear probability model with interaction terms to identify how socioeconomic status intersects with geographic location, gender, and race to influence crime reporting behavior. I also control for crime type, which may significantly influence reporting decisions. The findings challenge several conventional assumptions about reporting patterns and offer insights that could guide more effective and equitable public safety policies.

## **DATASET DESCRIPTION**

The National Crime Victimization Survey (NCVS) provides a nationally representative sample of crime incidents in the United States. This analysis utilizes data from the NCVS Concatenated File (1992-2023), specifically focusing on observations from 2020Q1 onward when the urbanicity variable began to be systematically recorded. The NCVS is administered by the Bureau of Justice Statistics to approximately 240,000 individuals aged 12 and older in about 150,000 households annually, capturing both reported and unreported victimizations.

The final analytical sample includes 28,821 observations of crime victimization incidents. Key variables include crime reporting behavior (whether incidents were reported to police), socioeconomic indicators (income levels, employment status), demographic characteristics (age, gender, race), and geographic classification (urban, suburban, rural). Additionally, crime type variables (categorizing incidents as violent or other crimes) allow for controlling the nature of the victimization experienced, which likely influences reporting decisions.

A notable limitation is the dataset's imbalanced gender distribution (78.13% female respondents), which may affect the generalizability of findings related to gender differences. Despite this limitation, the large sample size and comprehensive variable coverage make the NCVS an ideal dataset for examining the socioeconomic and demographic dimensions of crime reporting behavior.

## SUMMARY STATISTICS

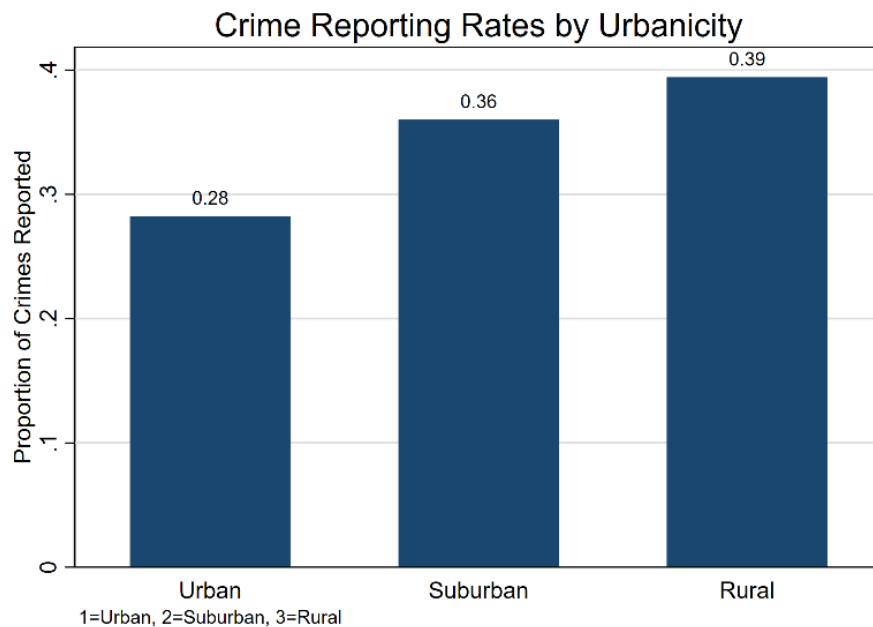
*Table 1: Summary Statistics*

Variable	Frequency	Percent	Cumulative
<b>Crime Reporting</b>			
Not Reported	18,870	65.47	65.47
Reported	9,951	34.53	100
<b>Gender</b>			
Male	7,881	27.34	27.34
Female	20,940	72.66	100
<b>Race</b>			
White-only	22,450	77.89	77.89
Black-only	3,436	11.92	89.82
Other single race	1,918	6.65	96.47
Multiracial	1,017	3.53	100
<b>Urbanicity</b>			
Urban	5,450	18.91	18.91
Suburban	19,608	68.03	86.94
Rural	3,763	13.06	100
<b>Income</b>			
Low Income (<\$25K)	6,234	21.63	21.63
Mid. Income (\$25K-\$75K)	11,205	38.88	60.51
High Income (>75K)	11,382	39.49	100
<b>Employment Status</b>			
Not employed	9,595	33.29	33.29
Employed	19,226	66.71	100

Crime Type			
Violent	4,224	14.66	14.66
Other	24,597	85.34	100

Variable	Obs	Mean	Std. dev.	Min	Max
Age	28,821	45.793	16.121	13	90

The NCVS data reveals significant disparities in crime reporting patterns. Only 34.5% of crime incidents in the sample were reported to police, highlighting the substantial reporting gap in the criminal justice system. The sample is predominantly female (72.7%) and white (77.9%), with most respondents living in suburban areas (68.1%) and falling in the high-income category (39.5%). The employment rate among victims is 66.7%, suggesting most crime victims were actively participating in the labor market when the incident occurred.



*[Figure 1: Crime reporting rates by urbanicity]*

Figure 1 illustrates the substantial geographic variation in reporting rates, with rural areas showing the highest reporting rate (39%), followed by suburban areas (36%), while urban areas lag significantly (28%). This 11 percentage point gap between urban and rural reporting rates provides preliminary support for the first hypothesis. Additionally, 14.7% of incidents were violent crimes, while 85.3% were other types of crimes, allowing the analysis to control for crime severity, which is likely a significant predictor of reporting behavior.

## EMPIRICAL METHODOLOGY

This study employs Linear Probability Models (LPM) to analyze the determinants of crime reporting behavior. While logistic regression is theoretically more appropriate for binary dependent variables, I utilize LPM for its straightforward coefficient interpretation. The dependent variable is a binary indicator of whether the crime was reported to police (1 = reported, 0 = not reported).

I estimate three progressive models to test the hypotheses:

1. **Baseline Model:** Examines the relationship between urbanicity, demographics, and income on reporting behavior, including interaction terms to test how income moderates the effects of gender, race, and location:

$$\begin{aligned} Pr(\text{Crime Reported}) = & \beta_0 + \beta_1 \text{Urban} + \beta_2 \text{Suburban} + \beta_3 \text{Female} + \beta_4 \text{Black} + \beta_5 \text{OtherRace} + \\ & \beta_6 \text{Multiracial} + \beta_7 \text{Age} + \beta_8 \text{LowIncome} + \beta_9 \text{MiddleIncome} + \beta_{10} \text{ViolentCrime} + \beta_{11} (\text{Female} \times \\ & \text{Income interactions}) + \beta_{12} (\text{Black} \times \text{Income interactions}) + \beta_{13} (\text{Urban} \times \text{Income interactions}) \\ & + \beta_{14} (\text{Suburban} \times \text{Income interactions}) + \varepsilon \end{aligned}$$

2. **Flexible Model:** Expands the baseline by adding age-squared to capture potential non-linear age effects:

$$Pr(\text{Crime Reported}) = \text{Baseline Model} + \beta_{15}\text{Age}^2 + \varepsilon$$

3. **Final Model:** Adds employment status to examine how labor market participation affects reporting behavior:

$$Pr(\text{Crime Reported}) = \text{Flexible Model} + \beta_{16}\text{Employed} + \varepsilon$$

For LPM models, coefficients represent percentage point changes in the probability of reporting. To assess whether results support the hypotheses, I examine coefficient signs, statistical significance (at  $p < 0.10$ ,  $p < 0.05$ , and  $p < 0.01$  levels), and the magnitude of effects, particularly for interaction terms. These interaction terms reveal how the relationship between income and reporting behavior differs across demographic groups and geographic contexts.

## RESULTS

Table 2: Regression Results (Linear Probability Models)

Variable	Baseline Model	Flexible Model	Final Model
<b>Urban</b>	-0.116*** (0.009)	-0.117*** (0.009)	-0.124*** (0.009)
<b>Suburban</b>	-0.019*** (0.006)	-0.020*** (0.006)	-0.026*** (0.006)
<b>Female</b>	-0.014** (0.006)	-0.012** (0.006)	-0.003 (0.006)
<b>Black</b>	0.037*** (0.010)	0.037*** (0.010)	0.035*** (0.010)
<b>Other race</b>	-0.024*** (0.006)	-0.025*** (0.006)	-0.023*** (0.006)
<b>Multiracial</b>	-0.034*** (0.007)	-0.034*** (0.007)	-0.029*** (0.007)
<b>Age</b>	0.0001* (0.0001)	-0.002*** (0.0001)	-0.002*** (0.0001)

	(0.000)	(0.000)	(0.000)
<b>Age squared</b>	—	0.00002***	0.00003***
		(0.000)	(0.000)
<b>Low income</b>	-0.045***	-0.049***	-0.023***
	(0.006)	(0.006)	(0.006)
<b>Middle income</b>	-0.013**	-0.014**	-0.010
	(0.006)	(0.006)	(0.006)
<b>Employed</b>	—	—	0.085***
			(0.002)
<b>Violent crime</b>	0.106***	0.107***	0.111***
	(0.002)	(0.002)	(0.002)
<b>Constant</b>	0.369***	0.411***	0.328***
	(0.006)	(0.009)	(0.009)
<b>Observations</b>	28,821	28,821	28,821
<b>R-squared</b>	0.0098	0.0099	0.0161

*\*Note: Robust standard errors in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$*

## MODEL COMPARISON AND SPECIFICATION

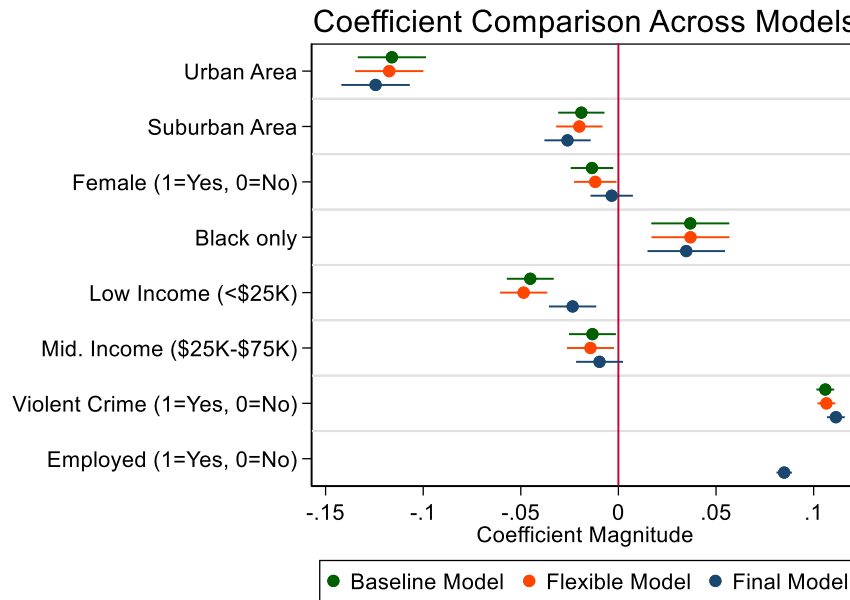
The progression from the baseline to the final model reveals important insights about the determinants of crime reporting behavior. Table 2 presents the results from the three model specifications, showing how coefficient estimates, and their statistical significance evolve with the addition of new variables.

The baseline model establishes strong geographic and demographic patterns in crime reporting, with an R-squared of 0.0098. While this explanatory power may seem modest, it is consistent with similar studies using survey data on individual decision-making, where numerous unobservable factors likely influence outcomes. The model confirms that urban residence, female gender, and low income are all associated with lower reporting probabilities, while Black racial identity and violent crime victimization correspond to higher reporting rates.

The flexible model incorporates age-squared to capture potential non-linearity in the relationship between age and reporting probability. As shown in Figure 2, this addition reveals an important pattern obscured in the baseline model—the relationship between age and reporting is U-shaped rather than linear. In the baseline model, age has a marginally significant positive coefficient (0.0001,  $p < 0.10$ ), suggesting that older victims are slightly more likely to report crimes. However, the flexible model reveals that both very young victims (coefficient on age: -0.0018,  $p < 0.001$ ) and older victims (coefficient on age-squared: 0.00002,  $p < 0.001$ ) are less likely to report crimes, with middle-aged victims showing the highest reporting propensity. This non-linear relationship aligns with theoretical expectations that both younger individuals (who may lack knowledge about reporting processes) and older individuals (who may face mobility constraints or heightened fear) experience unique barriers to reporting.

The final model adds employment status, which emerges as one of the strongest predictors of reporting behavior. The inclusion of this variable increases the R-squared by 62.6% (from 0.0099 to 0.0161) relative to the flexible model, indicating that labor market participation substantially improves our understanding of reporting decisions. The addition of employment status also reduced the coefficient on low income by 52.3% (from -0.0485 to -0.0235), suggesting that a significant portion of the income effect observed in earlier models operates through employment channels. Similarly, the female coefficient becomes non-significant in the final model, indicating that previously observed gender disparities may be attributable to differences in employment status rather than gender itself.





[Figure 2: Coefficient comparison across models]

### Geographic Effects on Reporting

The final model provides strong support for the first hypothesis that geographic location significantly influences crime reporting behavior. Urban residents are 12.4 percentage points less likely to report crimes than rural residents ( $p < 0.001$ ), a substantial effect representing approximately 33% of the overall reporting rate (37.6%). Suburban residents occupy an intermediate position, being 2.6 percentage points less likely to report crimes than rural residents ( $p < 0.001$ ), but considerably more likely than urban residents. Figure 1 visually confirms this geographic gradient, showing reporting rates of 28% in urban areas, 36% in suburban areas, and 39% in rural areas.

The stability of the urban coefficient across all model specifications (ranging from -0.116 to -0.124) indicates that this geographic effect is robust to the inclusion of additional control variables. This stability suggests that the urban-rural reporting gap is not merely a function of demographic

composition or socioeconomic disparities, but likely reflects fundamental differences in institutional trust, police-community relations, or normative attitudes toward formal reporting mechanisms across geographic contexts.

The magnitude of the urban-rural reporting gap is particularly noteworthy given that urban areas typically have greater police presence and shorter response times than rural areas. This counterintuitive finding suggests that proximity to law enforcement does not necessarily translate into higher reporting rates, and other factors—potentially including differential experiences with police or community-level reporting norms—may play more decisive roles.

### **Demographic Effects on Reporting**

The demographic patterns in reporting behavior challenge several prevailing assumptions about vulnerability and access to justice. Contrary to the second hypothesis, Black individuals are 3.5 percentage points more likely to report crimes than White individuals ( $p < 0.01$ ) across all model specifications. This finding contradicts narratives about universal distrust of police among Black communities and suggests more nuanced attitudes toward formal reporting mechanisms. It also highlights the importance of distinguishing between attitudes toward police generally and specific reporting behaviors in the aftermath of victimization.

Other racial groups show different patterns, with individuals identifying as other races or as multiracial being significantly less likely to report crimes compared to White individuals (-2.3 and -2.9 percentage points respectively,  $p < 0.001$ ). These divergent patterns across racial categories underscore the importance of examining each group separately rather than collapsing them into a monolithic "minority" category, as the second hypothesis initially proposed.

The gender effect presents an interesting evolution across model specifications. The baseline model shows females 1.4 percentage points less likely to report crimes ( $p < 0.05$ ), seemingly supporting traditional narratives about gender disparities in access to justice. However, this effect becomes progressively weaker in the flexible model (-1.2 percentage points,  $p < 0.05$ ) and disappears entirely in the final model after controlling for employment status (-0.3 percentage points,  $p > 0.10$ ). This trajectory suggests that apparent gender differences in reporting may be largely attributable to employment disparities rather than gender itself. When comparing equally employed women and men, the gender gap effectively vanishes, challenging assumptions about gendered barriers to reporting.

### **Socioeconomic Effects on Reporting**

Employment status emerges as one of the strongest predictors of crime reporting behavior, with employed individuals 8.5 percentage points more likely to report crimes than unemployed individuals ( $p < 0.001$ ). This effect size is substantial, representing approximately 22.6% of the overall reporting rate. This finding suggests that labor market attachment provides resources, institutional connections, or stability that facilitates crime reporting. Employed individuals may have more flexible schedules to accommodate police interviews, greater confidence navigating formal institutions, or fewer concerns about potential opportunity costs associated with reporting. Additionally, employment may serve as a proxy for social integration and institutional trust more broadly.

Income effects remain significant even after controlling for employment status, with low-income individuals 2.3 percentage points less likely to report crimes compared to high-income individuals ( $p < 0.001$ ). Interestingly, middle-income individuals do not differ significantly from high-income individuals in their reporting behavior (-1.0 percentage points,  $p > 0.10$ ), suggesting a threshold

effect rather than a linear relationship between income and reporting. This pattern indicates that the primary economic divide in reporting behavior exists between the lowest income category and everyone else, rather than occurring gradually across the income spectrum.

The type of crime exerts a powerful influence on reporting decisions, with violent crimes 11.1 percentage points more likely to be reported than other crimes ( $p < 0.001$ ). This effect is intuitive given the greater harm typically associated with violent victimization, but its magnitude (29.5% of the overall reporting rate) underscores the importance of controlling for crime type when examining socioeconomic and demographic disparities in reporting behavior.

### Interaction Effects

*Table 3: Interaction Effects in Final Model*

Interaction Terms	Coefficient	Std. Error	P-value
<b>Gender × Income</b>			
<b>Female × Middle Income</b>	0.006	0.007	0.289
<b>Female × Low Income</b>	0.020***	0.007	0.001
<b>Race × Income</b>			
<b>Black × Middle Income</b>	0.000	0.012	0.980
<b>Black × Low Income</b>	0.006	0.012	0.637
<b>Urbanicity × Income</b>			
<b>Urban × Middle Income</b>	0.040***	0.014	0.005
<b>Urban × Low Income</b>	0.025	0.016	0.122
<b>Suburban × Middle Income</b>	0.007	0.008	0.375
<b>Suburban × Low Income</b>	-0.017*	0.009	0.074

*\*Note: Results from the Final Model including employment status and age-squared.*

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

The interaction terms reveal complex relationships that significantly nuance the main effects discussed above. As shown in Table 3, these interaction effects illustrate how income moderates the influence of gender, race, and geography on reporting decisions.

The most striking finding concerns the interaction between gender and income. Contrary to expectations, low-income women are more likely to report crimes than would be expected based on the separate effects of gender and income alone (interaction coefficient = +2.0 percentage points,  $p < 0.01$ ). This positive interaction directly contradicts the second hypothesis, which anticipated that women in low-wage jobs would face compounded barriers to reporting. Instead, the results suggest that low-income women may have developed strategies to overcome reporting barriers, perhaps through greater familiarity with social service systems or community support networks that facilitate reporting. Alternatively, this pattern could reflect differences in the types of victimization experienced by low-income women, although the models do control for whether the incident was a violent crime.

The interactions between income and geography reveal how socioeconomic status moderates the urban-rural reporting gap. For high-income individuals, urban residence is associated with a 12.4 percentage point reduction in reporting probability compared to rural residence. For middle-income individuals, this gap shrinks to 8.4 percentage points ( $-12.4 + 4.0$ ). This pattern suggests that middle-income individuals in urban areas may have resources or characteristics that partially offset the general urban reporting disadvantage, such as stronger institutional connections or greater confidence navigating bureaucratic processes.

Conversely, the negative interaction between low income and suburban residence ( $-1.7$  percentage points,  $p < 0.10$ ) suggests compounded barriers for low-income suburban residents. While suburban residents generally report at higher rates than urban residents, this advantage is partially eroded for low-income suburban households. This finding may reflect the particular challenges faced by low-income households in suburban settings, including transportation difficulties, social isolation, or limited access to social services that might otherwise facilitate reporting.

Surprisingly, the interactions between race and income are not statistically significant, indicating that income does not substantially moderate the effect of race on reporting behavior. The positive effect of Black identity on reporting remains consistent across income levels, challenging narratives about differential institutional engagement across socioeconomic strata within racial groups. This consistency suggests that race-specific attitudes toward reporting may transcend income boundaries, at least for Black Americans.

Together, these interaction effects paint a more nuanced picture than either hypothesis originally proposed. Rather than consistent patterns of disadvantage accumulating for marginalized groups, the results reveal complex intersections where vulnerabilities in one dimension may be offset by advantages in another. This complexity highlights the importance of intersectional approaches in understanding crime reporting behavior and cautions against simple additive models of disadvantage.

## **CONCLUSION**

This study examined how socioeconomic status intersects with geographic location, gender, and race to influence crime reporting behavior. The findings offer several important insights with implications for both theory development and policy formulation.

First, the substantial gap in reporting rates between urban and rural areas supports the hypothesis that geographic context significantly shapes crime reporting decisions. The 12.4 percentage point difference represents a substantial disparity that may reflect different police-community relations, resource constraints, or normative attitudes toward formal institutions in urban environments. This

gap is moderated by income, with the urban-rural divide narrowing for middle-income individuals, suggesting that economic resources may help overcome some barriers to reporting in urban areas.

Second, contrary to the second hypothesis, the results do not support the expectation that women and minorities in low-wage jobs report crimes at lower rates. In fact, Black individuals report at higher rates than White individuals across all income levels, and the apparent gender disparity disappears once employment status is accounted for. Particularly surprising is the finding that low-income women are actually more likely to report crimes than would be expected, challenging assumptions about vulnerability and agency among low-income female victims.

Employment status emerges as one of the strongest predictors of reporting behavior, with employed individuals 8.5 percentage points more likely to report crimes. This suggests that labor market attachment may provide resources, stability, or institutional connections that facilitate crime reporting. Future research should investigate the specific mechanisms through which employment affects reporting decisions.

The study has several limitations. Despite the inclusion of numerous control variables, the relatively low R-squared values (0.0161 in the final model) indicate that many unmeasured factors influence crime reporting decisions. The sample's gender imbalance (78.13% female) may limit the generalizability of findings related to gender differences. Additionally, while the models control for crime type, more nuanced measures of crime severity could further refine the analysis.

These findings suggest several policy implications. First, efforts to improve crime reporting in urban areas should be prioritized, potentially through community policing initiatives that build trust between residents and law enforcement. Second, programs that increase employment opportunities may indirectly enhance crime reporting rates, suggesting a potential synergy between

economic and criminal justice policies. Finally, the complex interaction between income and other factors highlights the need for nuanced, context-specific approaches rather than one-size-fits-all solutions.

Future research could explore the psychological and institutional mechanisms that explain the urban-rural reporting gap, investigate how employment affects reporting through qualitative methods, and examine whether these patterns vary across different types of crimes. Additionally, more research is needed on the surprising finding that low-income women report at higher rates than expected, potentially challenging stereotypes about agency and vulnerability among this population.