If you’re a criminal defendant, it may help—a lot—to be a woman. At least, that’s what Professor Sonja Starr’s research on federal criminal cases suggests.

Starr’s recent paper, “Estimating Gender Disparities in Federal Criminal Cases,” looks closely at a large dataset of federal cases, and reveals some significant findings. After controlling for the arrest offense, criminal history, and other prior characteristics, “men receive 63 percent longer sentences on average than women do,” and “[w]omen are…twice as likely to avoid incarceration if convicted.” This gender gap is about six times as large as the racial disparity that Starr found in another recent paper, “Racial Disparity in Federal Criminal Charging and Its Sentencing Consequences.”

Other studies have shown gender disparity in criminal cases, but not as pronounced as Starr’s findings. This is because she is looking at “a larger swath of the criminal justice process” in her analysis, she says. The paper states, “Existing studies have typically focused on single stages of the criminal process in isolation”—in particular, the judge’s final sentencing decision. These studies compare actual sentencing outcomes after controlling for the recommended sentence associated with the defendant’s ultimate conviction.

The problem with this, Starr says, is that “the key control variable is itself the result of a host of discretionary decisions made earlier in the justice process”—including prosecutors’ charging and plea-bargaining decisions. Starr’s research incorporates disparities found at those earlier stages, and finds that “more disparity is introduced at each phase of the justice process.”

After estimating the amount of disparity left unexplained by the arrest offense and other control variables, the paper explores “why these gaps exist—and, in particular, whether unobserved differences between men and women might justify them.” Starr explores several potential mitigating factors, such as the “girlfriend theory” (that women “might be viewed as…mere accessories of their male romantic partners”), the role of women as primary caregivers to their children, and the “theory that female defendants receive leniency because they are more cooperative with the government.” Although each of these theories found some support in the data, none appeared capable of explaining anything close to the disparity that Starr found.

Starr emphasizes that it is not possible to “prove” gender discrimination with data like hers, because it is always possible that two seemingly similar cases could differ in ways not captured by the data. Given the size of the apparent gender gap and the richness of the dataset (which allowed many alternative explanations to be explored), however, Starr says that there is “pretty good reason to suspect that disparate treatment may be one of the causes of this gap.”

If men and women are being treated differently by prosecutors and judges, what should be done about it? Starr leaves that question to policymakers, but she does note that the solution “is not necessarily to lock up a lot more women, but perhaps to reconsider the decision-making criteria that are applied to men. About one in every 50 American men is currently behind bars, and we could think about gender disparity as perhaps being a key dimension of that problem.”