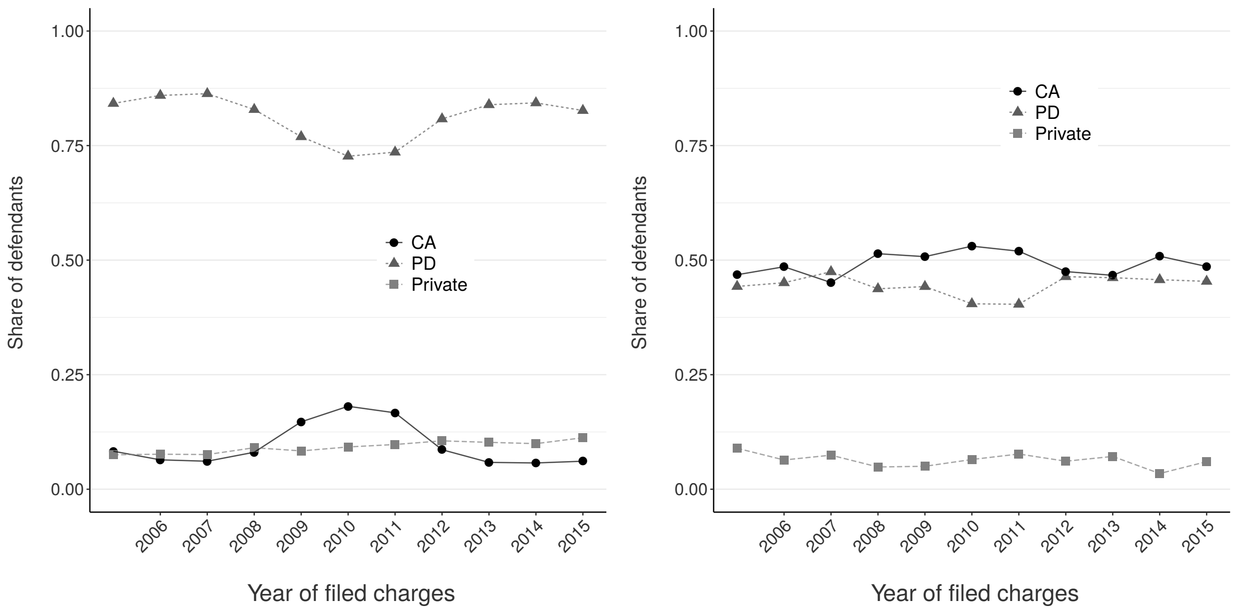
Online supplementary appendix

A Supplementary figures and tables

Figure A.1: San Francisco: The distribution of defendants across attorney types and over time, by filing year

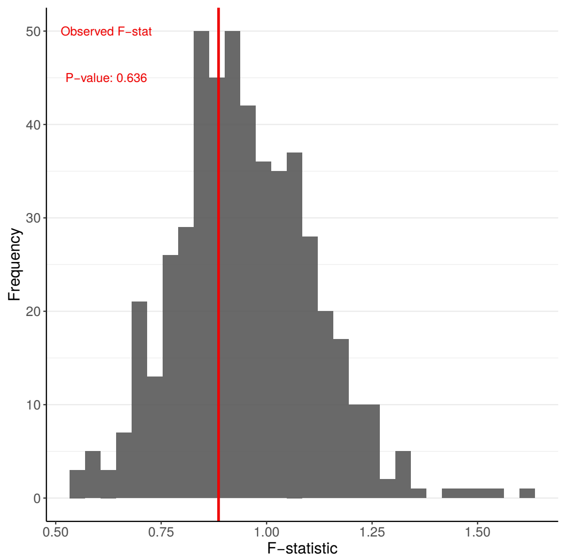


(a) Single defendant

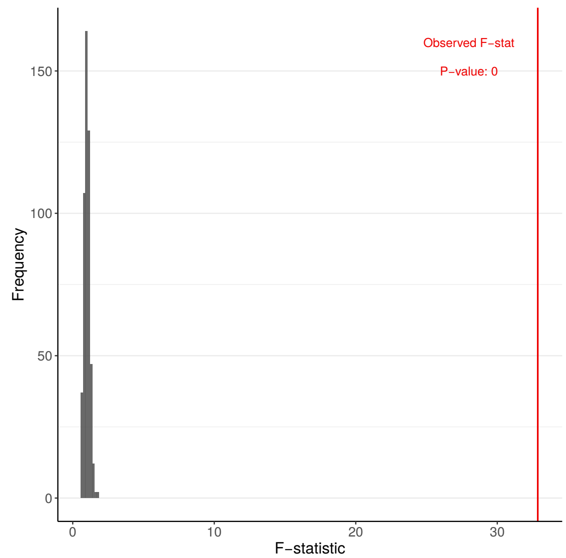
(b) Multiple defendants

Notes: The figure shows the distribution of criminal defendants in San Francisco across attorney types. The left plot shows the distribution of defendants across attorney types in cases with a single defendant. The right plot shows the distribution of defendants across attorney types in cases with multiple defendants.

Figure A.2: San Francisco: Monte-Carlo permutations of attorney type assignment within a case: F-statistic using Offense codes and controls



(a) Within multiple defendant case



(b) Across single defendant cases

Notes: Each one of the plots above uses Monte-Carlo simulations to asses whether the observed F-statistic is likely under a mechanism which randomly assigned defendants across attorney types. Fischman (2011), Abrams et al. (2012) and Abrams and Fackler (2017) all used similar Monte-Carlo simulation procedures when assessing covariate balance and to correct finite sample coverage concerns with the asymptotic distribution of the conventional F-statistic. The red line shows the observed F-statistic and the histogram plots an approximation of the distribution of the F-statistic under a random assignment mechanism using 1,000 random re-labellings of defendants across attorney types. I randomly permuted/shuffled which defendants have been assigned to a PD relative to a CA, and then estimated the F-statistic for the null that all the coefficients are equal to zero. In the multiple defendant sample the permutations are done within a case. The number of re-labellings we use is 1,000 and it is similar to what is commonly used in the statistics literature. For example, Athey et al. (2018) and Anderson and Magruder (2017) use 1,000 draws/re-labellings; and Keele and Miratrix (2019) use 500 draws/re-labellings.

Table A.1: San Francisco: Differences in observable characteristics between defendants who are assigned PD and CA

	All indigent	All Multiple	Multiple PD & CA		Co-defendant PD & CA	
	(1)	(2)	(3)	(4)	(5)	(6)
Age	1.625***	1.064***	0.630**	0.158	0.225	0.225
	(0.124)	(0.252)	(0.280)	(0.213)	(0.312)	(0.238)
Female	−0.050***	0.010	−0.002	−0.006	−0.007	−0.007
	(0.004)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)
White	0.084***	0.023**	0.005	0.003	−0.003	−0.003
	(0.005)	(0.010)	(0.012)	(0.008)	(0.013)	(0.009)
Black	−0.091***	−0.030***	−0.004	−0.006	−0.0003	−0.0003
	(0.005)	(0.011)	(0.012)	(0.008)	(0.013)	(0.009)
Hispanic	−0.010**	0.006	0.0002	−0.0005	0.005	0.005
	(0.004)	(0.008)	(0.010)	(0.009)	(0.011)	(0.010)
Felony	−0.220***	−0.069***	−0.007	0.001	0.001	0.001
	(0.005)	(0.008)	(0.008)	(0.002)	(0.009)	(0.002)
Predicted prison term	−0.764***	−0.782***	−0.171	0.007	0.090	0.090
	(0.090)	(0.161)	(0.177)	(0.131)	(0.174)	(0.143)
Predicted convicted	−0.018***	−0.009***	−0.0001	0.001	0.001	0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Predicted prison	−0.002***	−0.002***	0.00001	0.0001	0.0004	0.0004
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Most severe	–	0.024**	0.027**	0.010	0.016	0.016
	(0.003)	(0.011)	(0.012)	(0.016)	(0.013)	(0.018)
Num. prior incarceration	−0.053***	−0.008	−0.001	−0.010	−0.018	−0.018
	(0.007)	(0.011)	(0.013)	(0.012)	(0.015)	(0.014)
Num. prior convictions	−0.256***	−0.001	0.056**	0.037	0.025	0.025
	(0.015)	(0.025)	(0.028)	(0.024)	(0.032)	(0.028)
Num. prior incidents	−0.873***	0.053	0.193*	0.101	0.115	0.115
	(0.054)	(0.093)	(0.104)	(0.087)	(0.121)	(0.100)
Observations	67,620	8,975	7,164	7,164	5,826	5,826
Case FE	No	No	No	Yes	No	Yes

Notes: Each cell in the table contains the coefficient on an indicator whether the defendant was initially assigned a PD. The table reports the estimates of the β coefficient from model (1). Standard errors are clustered-robust at the case level. Columns 3 and 4 include all multiple defendant cases with both a PD and a CA within each case. Columns 5-6 include only multiple defendant cases with exactly two indigent defendants that one was assigned a PD and the other a CA. For example, a case with 3 indigent defendants that two of which are represented by CAs and the third by a PD will be included in columns 3 and 4 but not in columns 5 and 6. Notice also that in columns 5 and 6 the number of individuals within each case that are assigned to a PD is exactly the same as the number that is assigned to a CA. In this type of a balanced design the estimates in columns 5 and 6 are mechanically the same; however, the standard-errors are affected by the inclusion of case-level FEs in the regression specification. This mechanical equality between columns 5 and 6 in the point estimates would not have hold if continuous control variables would have also been included in the right hand side of the regression specification.

*p<0.1; **p<0.05; ***p<0.01

Table A.2: Variation in defendant characteristics within a multiple defendant case

	Multiple	Co-defendants
Obs.	2.143	2.000
Black & Non-Black	0.223	0.215
Hispanic & Non-Hispanic	0.333	0.314
White & Non-White	0.246	0.237
Black & White	0.224	0.218
Black & Hispanic	0.073	0.068
White & Hispanic	0.083	0.076
Felony & Non-Felony	0.014	0.012
Prior arrest & No prior arrest	0.338	0.322
Prior conv. & No prior conv.	0.156	0.158
Prior incar. & No prior incar.	0.287	0.280

Table A.3: San Francisco: The effect of having a PD vs. a CA on the case sentencing outcomes when controlling for attorney characteristics

	Initial PD effect			
	(1)	(2)	(3)	(4)
asinh(Prison term)	-0.118^{***}	-0.119^{***}	-0.101	-0.109^{*}
	(0.031)	(0.031)	(0.064)	(0.063)
Prison	-0.021^{***}	-0.021^{***}	-0.015	-0.016
	(0.006)	(0.006)	(0.011)	(0.011)
Convicted	-0.040^{***}	-0.046^{***}	-0.002	-0.010
	(0.010)	(0.010)	(0.018)	(0.018)
Case FE	Yes	Yes	Yes	Yes
Defendant controls	No	Yes	No	Yes
Attorney controls	No	No	Yes	Yes
Observations	6,703	6,703	6,703	6,703

Notes: Each cell in the table contains the coefficient on an indicator whether the defendant was initially assigned a PD or a CA. The standard errors are cluster-robust at the case level. Both incarceration and prison terms are measured in months. I approximate the Log(·) function using the asinh(·) function which is a common procedure when the outcome of interest is both skewed and has a mass at zero. The attorney characteristics include all the covariates in Table 3. The number of observations in this table is smaller than in Table 2, 6703 vs. 7164, since in some of the observations the attorney type was available but the attorney name was either not available or was partially listed.
*p<0.1; **p<0.05; ***p<0.01

Table A.4: San Francisco: Changes in attorney characteristics between first and terminating attorneys

	CA	PD	RE
Change attorney	0.207	0.521	0.161
Higher rank JD (US news)	0.010	0.080	0.006
Higher rank BA (US news)	0.079	0.232	0.060
Higher experience	0.163	0.354	0.142
Lower rank JD (US news)	0.011	0.089	0.014
Lower rank BA (US news)	0.091	0.214	0.070
Lower experience	0.100	0.167	0.080

B Covariate indices for charge severity measures

To quantify the gaps in the severity of the filed charges between defendants that are assigned a PD and those assigned a CA, I consider a simple summary measure of the selection based on a Oaxaca decomposition. A trial outcome (e.g., incarceration length) Y_{ig} can be modelled by projecting it on a set of pre-trial charge characteristics:

$$Y_{ig} = X_{ig}\beta_g + \nu_g, \quad \text{where} \quad g = \text{PD, CA} \tag{3}$$

The coefficient vector β_g has a causal interpretation under certain conditions (Fortin et al., 2011), and the fitted values $X_g\hat{\beta}_g$ are independent of $\hat{\nu}_g$ by construction. The average difference in the trial outcome, $\bar{Y}_{\text{PD}} - \bar{Y}_{\text{CA}}$, between attorney types can be written as (Oaxaca, 1973),

$$\bar{Y}_{\text{PD}} - \bar{Y}_{\text{CA}} = \hat{\beta}_{\text{CA}}\left(\bar{X}_{\text{PD}} - \bar{X}_{\text{CA}}\right) + \left(\hat{\beta}_{\text{PD}} - \hat{\beta}_{\text{CA}}\right)\bar{X}_{\text{PD}} \tag{4}$$

The first element in (4), $\hat{\beta}_{\text{CA}}\left(\bar{X}_{\text{PD}} - \bar{X}_{\text{CA}}\right)$, is the average difference in charge characteristics re-weighted by the effect of each characteristic on the trial outcome among defendants who are represented by a CA. This term represents selection on observables and will be zero in a standard balance test when:

$$\bar{X}_{\text{PD}} = \bar{X}_{\text{CA}} \tag{5}$$

One can summarize the imbalance in initial charge characteristics by estimating the difference in covariate indices $X_i'\hat{\beta}_{\text{CA}}$ that reduces the dimension of the covariate vector X_i to a single dimensional index. The idea of summarizing imbalance by the covariates' relationship to the outcome surface has been proposed in the past by several different procedures (Bowers and Hansen, 2009; Paetzold and Winner, 2016; Leacy and Stuart, 2014).

In San Francisco, I use the covariate index, $X_i'\hat{\beta}_{\text{PD}}$, which is based on estimating β using only defendants that have been assigned a PD. More specifically, I regress each case outcome on a vector of charge, case and defendant characteristics such as demographic characteristics, criminal history, charge severity (e.g., felony, misdemeanor). The main covariates are listed in Appendix Table A.1 and Figure 2. In addition, I use SC and BCS codes which are 2-digit and 3-digit classifications of offenses to broader categories.[13]

[13] The classification is done by the California Department of Justice, https://oag.ca.gov/law/code-tables.