

Supplemental Appendix

Connections During Democratic Transitions: Insights from the Political Purge in Post-WWII France

Toke S. Aidt^{*}, Jean Lacroix[†], and Pierre-Guillaume Méon[‡]

^{*}tsa23@econ.cam.ac.uk, Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, United Kingdom and CESifo Munich.

[†]jean.lacroix@universite-paris-saclay.fr, Université Paris-Saclay, Faculté Jean Monnet, RITM, 54 boulevard Desgranges, 92330 Sceaux and CESifo Munich.

[‡]p-guillaume.meon@ulb.be, CEBRIG and Dulbéa, Université libre de Bruxelles (ULB), CP-114/03, Av. F.D., Roosevelt, 50, 1050, Brussels and CESifo Munich.

A A theory of direct and indirect connections

This appendix develops a theory of how court decisions can be influence via connections. The theory is framed to match the procedure used in France at the end of World War II to judge the politicians who had voted for the enabling act. Accordingly, we consider two courts that have to decide if a ban on political participation already imposed on the defendants stands or is overturned. Each court has its own standard of proof required for acquittal. The defendants or their supporters can report mitigating circumstances with the aim of influencing the court decisions. Reporting mitigating circumstances is cheap talk, so there is no guarantee that it will, in fact, influence the courts. The fundamental problem is that the defendants and their supporters have an incentive to claim mitigating circumstances whether there are such circumstances or not. We model two mechanisms that can overcome this problem. Both mechanisms are related to connections between the courts, on the one hand, and the defendants and their supporters, on the other. The first mechanism is *direct* connections between a defendant and a court. The second is *indirect* connections operating via a third party (a supporter of the defendant). We show that both mechanisms can under certain circumstances help defendants with connections get acquitted by the court to which they are connected.

A.1 Courts decisions

Two courts, indexed by $c \in \{L, U\}$, are tasked with judging a fixed number of defendants indexed by $i \in D$. Before any evidence is heard, the two courts set their own bar for acquittal. After that, evidence is presented to the courts (related to whether the defendant had been sufficiently involved in the resistance), they receive letters from the defendants and/or their supporters claiming mitigating circumstances and make their decisions. Let x summarize the evidence presented to a court with $x \in (-\infty, \infty)$. A defendant who presents evidence stronger than the bar set by that court will be acquitted. To set the bar, we assume, as in Alesina and La Ferrara (2014), that the court's aim is to avoid making type I (convicting innocent defendants) and type II (not convicting guilty defendants) errors.¹ The weights that court c puts on type I and II errors are α_c and $1 - \alpha_c$, respectively. The evidence presented to the courts is drawn from the cumulative distribution function $A_G(x)$ if the defendant is guilty (did not participate in the resistance) and from $A_I(x)$ if innocent (participated in the resistance) and the corresponding density functions are a_G and a_I . We assume that the defendants can be

¹This is consistent with the type of democratic purge that we study. A purge in an authoritarian regime would likely give little weight to type I errors, if any.

divided into subgroups based on fixed observable characteristics, such as profession, political affiliation, region of residence, age, religion, and other observable characteristics, and that the proportion of guilty defendants in those subgroups may be perceived by the two courts to be different. We let $g \in \{1, 2, \dots, N\}$ with $\cup_g D_g = D$ index these subgroups. The two courts assume that the proportion of guilty among defendants belonging to subgroup g is π_g . As a consequence, each court sets N different bars – one for each group g . The objective function of court c can, then, be written as

$$\min_{x(c,g)} \sum_{g=1}^N \alpha_c (1 - \pi_g) A_I(x(c, g)) + (1 - \alpha_c) \pi_g (1 - A_G(x(c, g))). \quad (1)$$

Calculating the first order conditions, the optimal bar for court c for defendants belonging to sub-group g is the solution to

$$\frac{\alpha_c}{1 - \alpha_c} \frac{1 - \pi_g}{\pi_g} = \frac{a_G(x^*(c, g))}{a_I(x^*(c, g))} \quad (2)$$

and denoted $x^*(c, g)$. Given the bar, the probability that a defendant belonging to group D_g is acquitted by court c is

$$Pr[c, g] = \pi_g (1 - A_G(x^*(c, g))) + (1 - \pi_g) (1 - A_I(x^*(c, g))) \quad (3)$$

$$\equiv 1 - H_g(x^*(c, g)), \quad (4)$$

where $H_g = \pi_g A_G + (1 - \pi_g) A_I$. After applying a first order linear approximation, we can write the probability of acquittal before court c of a defendant from subgroup g as

$$Pr[c, g] \approx a + b_c + b_g. \quad (5)$$

We observe that a defendant's chances of acquittal differ before the two courts for two reasons. First, the courts may weigh the risk of the two types of mistakes differently (differences in α_c). This would lead to systematic differences in the acquittal rates between the courts and is captured by b_c in the linear approximation. This represents differences in the objectives and procedures of the two courts that can lead to differences in acquittal rates. Second, the two courts may perceive, based on observable characteristics, some groups of defendants to be more likely to be guilty than others (due to differences in π_g). This is a manifestation of statistical discrimination and can explain systematic differences in acquittal rates between different sub-groups of defendants and is captured by b_g in equation (5).

A.2 Connections to the Courts and Mitigating Circumstances

For each defendant i , we assume that there may or may not be mitigating circumstances that he or one of his supporters can present to the courts with the aim of influencing the court decisions. We model this by θ which can take two values: θ_Y if there are mitigating circumstances and θ_N if not with $\theta_Y > \theta_N$. This is private information to the defendants and their supporters and not known to the courts and cannot be externally verified. Mitigating circumstances are communicated to the courts via letters. The content of a letter sent by defendant i either personally or via a supporter to court c is $l_{i,c} \in \{\theta_Y, \theta_N\}$. If a court c accepts a letter claiming that $\theta = \theta_Y$ for a defendant i belonging to subgroup g , then it is more likely that the defendant is acquitted and his acquittal probability increases by $\eta_{c,g} > 0$ and, using the linear approximation to the acquittal probability in equation (5), becomes

$$Pr[c, g] = a + b_c + b_g + \eta_{c,g}. \quad (6)$$

We can interpret a letter $l_{i,c} = \theta_N$ as not sending a letter containing arguments about mitigating circumstances to court c . Court c wants to base its judgment on all the facts and it needs to decide if mitigating circumstances should be taken into account or not. Formally, the objective of court c in relation to the case of defendant i is to avoid mistakenly taking mitigating circumstances into account: $U_{i,c} = -(m_{i,c} - \theta)^2$ where m_i is interpreted as a decision to take mitigating circumstances into account ($m_{i,c} = \theta_Y$) or not ($m_{i,c} = \theta_N$) for defendant i . The optimal decision is $m_{i,c} = E(\theta|l_{i,c})$ where E is the expectation operator. If a court gets no letter for a defendant i , then its prior is that there are no mitigating circumstances and $m_{i,c} = \theta_N$. All the defendants want mitigating circumstances, if any, to be taken into account by the courts, but also to be acquitted regardless. The objective of defendant i with information θ , therefore, is $U_{i,c} = -(m_{i,c} - \theta - \gamma)^2$, where $\gamma > 0$ captures the desire to be acquitted regardless, i.e., the optimal decision by court c from the point of view of defendant i , $m_{i,c} = \theta + \gamma$, is biased in favor of taking mitigating circumstances into account whether there are, in fact, such circumstances. We model the interaction between a defendant i (or a supporter of defendant i , respectively) and each court c as a sequential game of asymmetric information where the defendant, firstly, after privately learning if there are mitigating circumstances, sends a letter to the court which, secondly, updates its beliefs about whether there are mitigating circumstances based on the content using Bayes rule when possible. The equilibrium concept is perfect Bayesian equilibrium. We assume that the motive to falsely claim mitigating circumstances is so strong that it is not possible, in general, for any defendant to write a letter that credibly claims mitigating circumstances ($\gamma > \frac{\theta_Y - \theta_N}{2} \equiv \bar{\gamma}$), i.e., we rule out informative perfect Bayesian equilibria of the type considered by Crawford

and Sobel (1982). In order to influence a court, a defendant needs to use his direct or indirect connections to that court.

A.2.1 Direct connections

All defendants can, in principle, communicate with the courts but a defendant with a *direct* connection to a court has an advantage in doing so. To model direct connections, we assume that communication is associated with a fixed cost.² Each defendant i faces a fixed cost, $f_{i,c}$, to connect to court c . A direct connection lowers the cost of communication. There can be many reasons for this, ranging from personal or professional ties to shared knowledge about the proper etiquette for communication in the context. The first proposition shows when and how direct connections can help a defendant get acquitted before court c .

Proposition 1. *There exist two values $\bar{f}$ and $\underline{f}$ with $\bar{f} > \underline{f}$ such that*

1. **Ineffective connected defendants:** Defendants with a cost of communicating $f_{i,c} < \underline{f}$ cannot convince court c and will not submit a letter claiming mitigating circumstances to that court.
2. **Effective connected defendants:** Defendants with a cost of communicating $\underline{f} \leq f_{i,c} \leq \bar{f}$ will submit a letter claiming mitigating circumstances ($l_{i,c} = \theta_Y$) to court c if and only if there are such circumstances and the court will believe the claim.
3. **Unconnected defendants:** Defendants with a cost of communicating $f_{i,c} > \bar{f}$ will not submit a letter claiming mitigating circumstances to court c even though the court would believe such a letter.

Proof. Consider a given court c . We need to find the values of the fixed cost such that 1) a defendant wants to communicate that $l_{i,c} = \theta_Y$ when that is, in fact, the case and 2) a defendant do not want to claim that $l_{i,c} = \theta_Y$ when $\theta = \theta_N$. Assume that the court believes $l_{i,c} = \theta_Y$ and sets $m_{i,c} = \theta_Y$ in response to a letter with such a claim and consider the game between defendant i and court c . First, suppose that $\theta = \theta_Y$, i.e., there are mitigating circumstances for defendant i . Let the cost of communicating with court c for defendant i be $f_{i,c}$. He will then get $-(\theta_Y - \theta_Y - \gamma)^2 - f_{i,c}$ if he sends the letter $l_{i,c} = \theta_Y$ and $-(\theta_N - \theta_Y - \gamma)^2$ if he does not send a letter. Comparing these payoffs, we see that it is in the interest of defendant i to send the letter $l_{i,c} = \theta_Y$ if

$$f_{i,c} \leq (\theta_Y - \theta_N)(2\gamma + (\theta_Y - \theta_N)) \equiv \bar{f}. \quad (7)$$

²See Grossman and Helpman (2002) for a similar approach to lobbying.

Second, suppose that $\theta = \theta_N$, i.e., there are no mitigating circumstances for defendant i . If defendant i has cost $f_{i,c}$, then he will get $-(\theta_Y - \theta_N - \gamma)^2 - f_{i,c}$ if he sends the letter $l_{i,c} = \theta_Y$ and $-(\theta_N - \theta_N - \gamma)^2$ if he does not send a letter. Comparing these payoffs, we see that it is in the interest of defendant i not to lie and send a letter with $l_{i,c} = \theta_Y$ when $\theta = \theta_N$ if

$$f_{i,c} \geq (\theta_Y - \theta_N)(2\gamma - (\theta_Y - \theta_N)) \equiv \underline{f}. \quad (8)$$

Clearly, $\bar{f} > \underline{f}$. Given these strategies, the court will update via Bayes Rule its belief to “mitigating circumstances” if and only if $f_{i,c} \in [\underline{f}, \bar{f}]$. The three cases in the proposition follows immediately from this. $\square$

The proposition shows that the defendants for each court are endogenously sorted into two subsets: the set of connected defendants, denoted K_c^d , and the set of unconnected defendants, denoted k_c^d , where the superscript d refers to direct connections. The set K_c^d of connected defendants consists of two types of defendants. While all defendants connected to court c have relatively low cost of communicating with that court ($f_{i,c} \leq \bar{f}$), only some of them are successful at convincing the court that mitigating circumstances are relevant. If they are “too” connected to the court ($f_{i,c} < \underline{f}$), then the court will not trust their letters. Indeed, in these cases, the cost of sending information is so low that defendants will do it regardless of whether there are mitigating circumstances or not. As a consequence, receiving a letter from them is uninformative. The defendants who can influence the court are those with a moderately good connection to it ($f_{i,c} \in [\underline{f}, \bar{f}]$). The court will believe what they have to say and so this group will submit letters claiming mitigating circumstances. For the set of unconnected defendants, the cost of communication is too high ($f_{i,c} > \bar{f}$) and they do not submit letters claiming mitigating circumstances to court c . This captures the first channel through which connections can influence court decisions. The objective of the court is to minimize judgement errors and it knows that defendants have an incentive to use their connection to the court to curry favors. This implies that defendants with “too good” a connection will not be trusted and it is only those for whom using their direct connection to the court is moderately costly who will be able to convey information about mitigating circumstances credibly and to allow the court to screen the information they receive directly from the defendant effectively. The proposition has the following empirical implication: if the population of defendants contains directly connected defendants as well as unconnected ones, then, on average, the connected defendants are at least as likely as those who are unconnected to that court to be acquitted *by the court to which they are connected* and strictly more likely if they are not all “too” connected.

A.2.2 Indirect connections

A defendant may also have an *indirect* connection to court c through a third party (T) who can submit *letters of support* to that court on behalf of the defendant. The third party supporters observe θ and, thus, know whether there are mitigating circumstances or not: they have already screened the information about mitigating circumstance and established whether it is credibility or not. Unlike the defendants, all third parties can send letters at low cost (for simplicity, we set the cost at zero for both courts, i.e., $f_c^T = 0$ for all c). This means that they are all “too” connected and the reason why third party supporters can influence the courts is, therefore, fundamentally different from the mechanism behind proposition 1. The reason that supporters of a defendant may be able to influence court c is that while they care about the defendant they support (U_i), they also, to some extent, align with the objective that court c uses to judge mitigating circumstances ($U_{i,c}$). Specifically, we assume that the third party supporters of defendant i have the following objective in relation to court c :

$$U_{i,c}^T = \beta_{i,c}U_{i,c} + (1 - \beta_{i,c})U_i = -\beta_{i,c}(m_{i,c} - \theta)^2 - (1 - \beta_{i,c})(m_{i,c} - \theta - \gamma)^2 \quad (9)$$

where $\beta_{i,c}$ determines the relative weight on the two considerations which may vary for defendants across the two courts. From the point of view of the third party, the optimal decision of court c for defendant i is $d_{i,c}^T = \theta + (1 - \beta_{i,c})\gamma$. One interpretation of this is that the judges may trust more people they know better – in part because these people would lose more if the evidence they provided proved to be wrong. The next proposition shows how indirect connections to a court can help a defendant to get acquitted.

Proposition 2. *There exists a $\bar{\beta} \in (0, 1)$ such that a third party with $\beta_{i,c} > \bar{\beta}$ who submits a letter on behalf of defendant i to court c can convince that court that there are mitigating circumstances for defendant i when that is the case. A letter to court c from a third party with $\beta_{i,c} \leq \bar{\beta}$ is not believed by that court.*

Proof. This is a standard cheap talk game. Consider court c and suppose that it believes the third party if a letter saying $l_{i,c} = \theta_Y$ is received. There is no reason not to send this letter if $\theta = \theta_Y$. Suppose, therefore, that $\theta = \theta_N$. In this case, the third party has an incentive to lie and write in the letter that $l_{i,c} = \theta_Y$. If he does write this, his payoff is $-\beta_{i,c}(\theta_Y - \theta_N)^2 - (1 - \beta_{i,c})(\theta_Y - \theta_N - \gamma)^2$. If he instead writes $l_{i,c} = \theta_N$, then his payoff is $-\beta_{i,c}(\theta_N - \theta_N)^2 - (1 - \beta_{i,c})(\theta_N - \theta_N - \gamma)^2 = (1 - \beta_{i,c})\gamma^2$. Comparing these two payoffs, we find that the third party will not be tempted to write a letter saying $l_{i,c} = \theta_Y$ when $\theta = \theta_N$ if

$$\beta_{i,c} > 1 - \frac{\theta_Y - \theta_N}{2\gamma} \equiv \bar{\beta}. \quad (10)$$

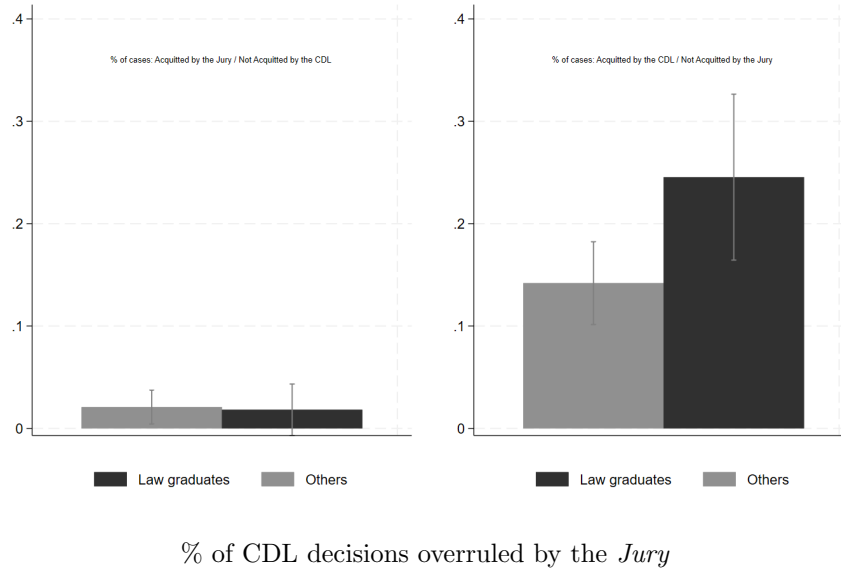
Clearly $\bar{\beta} < 1$. The cut-off $\bar{\beta} > 0$ because we assume that $\gamma > \frac{\theta_Y - \theta_N}{2} \equiv \bar{\gamma}$ for all i . Knowing this, court c will believe a letter of support claiming mitigating circumstances for defendant i coming from a third party with $\beta_{i,c} > \bar{\beta}$ and not otherwise. $\square$

The proposition says that third parties are effective at intervening on behalf of a defendant before court c if they at least to some degree share the same objective as the court. This splits the set of defendants into two subsets: those with indirect connections K_c^{id} and those without k_c^{id} , where superscript id refers to indirect connections. A letter from a third party who mostly cares about the defendant will not, in general, influence a court's decision. The advantage of having a third party intervening on behalf of a defendant is that third parties (to varying degrees) are concerned about the integrity of the legal process or by their own reputation and thus have objectives that partly overlap with those of the court. This is what makes their letters of support credible and influential. The connection channel is novel and requires that third-party supporters are able to screen the information about the defendant, communicate easily with the court, and align to some degree with the court's objective of minimizing judgment errors. While the first two features are necessary, on their own they are not sufficient. It is the third feature that enables indirect connections to work: when the court understands that the third party sufficiently shares its objective, it can benefit from the screening of information done by the third party and base its judgment on the information received through that channel.

Two empirical predictions flow from this. First, defendants supported by third parties trusted by a court are more likely to be acquitted than those who are not. Second, these defendants have more letters of support in their case files from third parties with a connection to that court than others.

B Baseline results - Additional Figures

Figure B.1: Percentage of cases overruled



C Baseline results - Robustness checks

C.1 Law graduates during purges: A summary

Table C.1: Share of Law Graduates throughout the purging process.

	Voted Yes in 1940	Faced Jury	Acquitted
Law Graduates	167	110	34
Total	569	399	100
Share	29.0%	27.5%	34%

C.2 Other Estimators

Table C.2: Connections of Law Graduates before the *Jury*: Difference-in-differences estimates

	(C.2.1)	(C.2.2)
	Dependent Variable: $Acquit_{i,c}$	
	Logit	Probit
Jury $\times$ LG	1.2 (.408)	2.15 (.885)
Observations	441	441
Pseudo R ²	.416	.414
<u>Controls:</u>		
Resistance and collaboration	Yes	Yes
Age and Religion	Yes	Yes
Political mandates	Yes	Yes
Journalist	Yes	Yes
<i>Département</i> fixed effect $\times$ Jury	Yes	Yes
<i>Département</i> fixed effect	Yes	Yes

Note: Standard errors are clustered at the *département* level (in parentheses). *Jury* is a dummy variable equal to one if the judgment was before the *Jury* and zero otherwise. LG is a dummy variable equal to one if the defendant was a Law graduate and zero otherwise. This table presents estimates of Equation (1). The controls are: Age and religion (Age, Jewishness); Journalist; Political mandates (Mayor, Special Role in the Assembly, parliamentarian of an occupied territory); Resistance and collaboration (Civilian resistance, Military resistance, Arrested by the Vichy regime, Mayor under *Vichy*). Each control is also interacted with the *Jury* dummy variable.

Table C.3: Coefficients of control variables included in the main estimation.

	(C.3.1)	(C.3.2)	(C.3.3)	(C.3.4)	(C.3.5)
Jury $\times$ LG	.106 (.0479)	.108 (.0418)	.0996 (.0418)	.14 (.0454)	.12 (.0515)
Jury $\times$ Civilian Resistance		.158 (.0344)	.166 (.0339)	.19 (.038)	.186 (.0389)
Jury $\times$ Arrested by Vichy		.219 (.0969)	.226 (.0952)	.253 (.105)	.252 (.108)
Jury $\times$ Military Resistance		.282 (.0729)	.257 (.0721)	.228 (.0809)	.227 (.0825)
Jury $\times$ Mayor under Vichy		-.0613 (.0404)	-.0108 (.0466)	.0368 (.0569)	.0358 (.0574)
Jury $\times$ Occupied Territory			-.0458 (.0349)	.111 (.0838)	.0771 (.0909)
Jury $\times$ Age			-.00222 (.00173)	-.00256 (.0021)	-.00297 (.00202)
Jury $\times$ Pre-War Mayor			-.0727 (.0459)	-.0138 (.0589)	-.0171 (.0585)
Jury $\times$ Journalist			.119 (.0509)	.0966 (.0591)	.0889 (.0604)
Jury $\times$ Jewish			.381 (.158)	.247 (.217)	.244 (.222)
Jury $\times$ Special Role Assembly			-.0589 (.0501)	.0135 (.0638)	.012 (.0638)
Jury $\times$ Workers					-.0448 (.0869)
Jury $\times$ Industrialists					-.00526 (.0651)
Jury $\times$ Agriculture					-.0421 (.063)
Jury $\times$ Teachers					-.0631 (.0673)
Observations	798	798	798	790	790
Adj. R ²	.0404	.271	.281	.292	.296
<i>Département</i> fixed effects	No	No	No	Yes	Yes
<i>Département</i> fixed effects $\times$ Jury	No	No	No	Yes	Yes

Note: Standard errors are clustered at the *département* level (in parentheses) *Jury* is a dummy variable equal to one if the judgment was before the *Jury* and zero otherwise. LG is a dummy variable equal to one if the defendant was a Law graduate and zero otherwise. This table presents estimates of the coefficients attached to control variables interacted with the *Jury* variable in Equation (1). The reduction in the number of observations in Columns 3.4 and 3.5 is due to the fact that within some *départements*, there is no variation in sentencing. The *département* fixed effect is, therefore, perfectly collinear with the outcome in those cases and eight observations are dropped.

C.3 Controls included one by one

Table C.4: Difference-in-differences estimates: including control variables one by one.

	(C.4.1)	(C.4.2)	(C.4.3)	(C.4.4)	(C.4.5)
	Dependent variable: $Acquit_{i,c}$				
Estimator	OLS	OLS	OLS	OLS	
Jury	.419 (.12)	.106 (.0231)	.218 (.0397)	-.0224 (.0243)	0 (.)
LG	-.0264 (.0301)	-.0276 (.0311)	-.0273 (.031)	-.0249 (.03)	-.0322 (.0343)
Jury $\times$ LG	.0978 (.0462)	.11 (.048)	.0958 (.0496)	.108 (.0418)	.13 (.0533)
Observations	798	798	798	798	790
Adj R ²	.0521	.0405	.049	.271	.105
<u>Controls:</u>					
Age and Religion	Yes				
Journalist		Yes			
Political mandates			Yes		
Resistance and collaboration WWII				Yes	
<i>Département</i> fixed effects $\times$ Jury					Yes
<i>Département</i> fixed effects					Yes

Note: *Jury* is a dummy variable equal to one if the judgment was made by the *Jury* and zero if it was made by a CDL. LG is a dummy variable equal to one if the defendant was a Law graduate and zero otherwise. This table presents estimates of Equation (1) but with the controls entered one by one. Controls are: Age and religion (Age, Jewishness); Journalist; Political mandates (Mayor, Special Role in the Assembly, parliamentarian of an occupied territory); Resistance and collaboration WWII (Civilian Resistance, Military resistance, Arrested by the Vichy regime, Mayor under *Vichy*). Each control variable is also interacted with the *Jury* dummy variable. Standard errors are clustered at the *département* level (in parentheses)

C.4 Counterfactual analysis

C.4.1 The method

We model the decision by the *Jury* to clear a defendant as a function of variables capturing participation of the defendant in the civilian and military resistance, which was the criteria officially used by the *Jury* to decide each case, controlling for the decision by CDLs. Specifically, we estimate the following regression on the sample of decisions of the *Jury*:

$$Acquitted_{Jury,i} = \alpha + \beta_1 Acquitted_{CDL,i} + \beta_2 CivilianResistance_i + \beta_3 MilitaryResistance_i + \varepsilon_i.$$

We then use the estimates of $\beta_1, \beta_2, \beta_3$ ($\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$) to compute each defendant's individual probability to be acquitted given the information retrieved from his dossier and the decision by the CDL that considered case. The estimated probability or rate given by $Counterfactual_{Jury,i} = \hat{\beta}_1 Acquitted_{CDL,i} + \hat{\beta}_2 CivilianResistance_i + \hat{\beta}_3 MilitaryResistance_i$ provides a counterfactual benchmark based on the official criteria used by the *Jury* against which to compare actual acquittal rates. We average these counterfactual acquittal rates at the group level.

C.4.2 Comparison between counterfactual and actual acquittal rates

Table C.5: Counterfactual versus actual acquittal rates for non-Law graduates and Law graduates

	Non-Law Graduates	Law graduates
Counterfactual	23.7%	22.4%
Actual	22.8%	30.9%

C.4.3 Placebo tests and control variables

Table C.6: Difference-in-differences estimates for different groups (Part I)

[illegible]

Table C.6 - Difference-in-differences estimates for different groups(PART II)

		CDL			Jury			Diff-in-Diff	
		(C.6.1)	(C.6.2)	(C.6.3)	(C.6.4)	(C.6.5)	(C.6.6)	(C.6.7)	
		Treated	Control	Diff=0	Treated	Control	Diff=0	$\Delta\Delta$	p-value
	<u>Treated:</u>	Group	Group	(p-value)	Group	Group	(p-value)		
Placebo Tests		Politics and political mandates							
	Senator	0.10	0.10	0.95	0.25	0.25	0.92	0.003	0.95
	Rightwing	0.11	0.09	0.42	0.25	0.25	0.93	-0.03	0.42
	Center	0.06	0.11	0.15	0.21	0.26	0.35	0.003	0.95
	Dynastic Politicians	0.06	0.11	0.29	0.27	0.25	0.70	0.07	0.27
		War experience							
	WWI veteran	0.11	0.09	0.63	0.25	0.25	0.89	-0.008	0.82
	WWII fighter	0.08	0.10	0.78	0.38	0.24	0.15	0.15	0.12
		Networks, clubs and religion							
	Free Masons	0.07	0.10	0.66	0.33	0.25	0.45	0.12	0.24
	Labour Unions	0.03	0.11	0.19	0.19	0.26	0.45	0.01	0.85
	Agr organization	0.13	0.10	0.50	0.26	0.25	0.85	-0.02	0.75
	War Medal	0.11	0.10	0.79	0.24	0.26	0.62	-0.03	0.46
	Légion d'Honneur	0.12	0.09	0.21	0.25	0.25	0.88	-0.03	0.40
	Veterans club	0.11	0.10	0.94	0.32	0.25	0.50	0.06	0.53
		Occupations							
	Civil servant	0.08	0.10	0.73	0.32	0.25	0.41	0.10	0.27
	Workers	0.11	0.10	0.90	0.25	0.25	0.99	-0.01	0.90
		Informational cues							
	Excluded by his party	0.12	0.09	0.37	0.31	0.23	0.09	0.05	0.28
	Signed Bergery motion	0.06	0.11	0.29	0.22	0.26	0.54	0.007	0.89
	National Mandate							0.00	0.80
	Conseiller général							-0.001	0.48
	Study Years							0.003	0.70

Note: The table reports mean comparisons and difference-in-differences estimates for different treatment and control groups. In part I, we select, one at the time, as the treatment group, the defendants belong to the groups defined by the control variables including in equation (1). In part II, we define the treatment group along a number of other dimensions (as a type of placebo test). Column C.6.1 presents the average acquittal rate of the treated group (defined in the left column) in front of the CDL (Comité Départementaux de Libération) whereas Column C.6.2 presents the average acquittal rate of the control group (i.e., all individuals not in the treated group) in front of the CDL. Column C.6.3 displays the difference between these two means. Column C.6.4 presents the average acquittal rate of the treated group in front of the *Jury* whereas Column C.6.5 presents the average acquittal rate of the control group (i.e., all individuals not in the treated group) in front of the *Jury*. Column C.6.6 displays the difference between Column C.6.4 and C.6.5. Column C.6.7 reports the difference-in-differences estimate of the difference in acquittal between a specific group (defined in the left column) and others in front of the *Jury* from a regression similar to equation (1) (but without any controls). This estimate is also by construction equal to the difference between the difference estimate in column C.6.6 and C.6.3. Standard errors are clustered at the département level.

C.5 Court-level estimates

Table C.7: The acquittal rates for Law graduates by court

	(C.7.1)	(C.7.2)
Dependent variable	$Acquit_i$	$Acquit_i$
Samples	CDLs	<i>Jury</i>
LG	-.0391 (.0333)	.101 (.0369)
Civilian Resistance	.0173 (.0285)	.208 (.0423)
Military Resistance	.19 (.0482)	.417 (.0765)
Arrested Vichy	-.00646 (.0982)	.247 (.0983)
Vichy Mayor	-.106 (.0418)	-.0696 (.0489)
Observations	395	395
Adj R ²	.126	.33
Age and Religion	Yes	Yes
Journalist	Yes	Yes
Political mandates	Yes	Yes
<i>Département</i> fixed effects	Yes	Yes

Note: Column C.6.1 reports the results from a linear probability model where the acquittal probability before the CDLs is regressed on the Law graduate dummy variable and all baseline control variables. Column C.6.2 reports the corresponding estimates for the *Jury*. Standard errors are clustered at the *département* level. Standard errors in parentheses.

C.6 Controlling for the composition of resistance in each department

Table C.8: Controlling for the composition of resistance in each department

Dep variable	(C.8.1) <i>Acquit_{i,c}</i>	(C.8.2) <i>Acquit_{i,c}</i>
Jury	-.11 (.0679)	.0763 (.122)
LG	-.0226 (.0296)	-.0294 (.0296)
Jury $\times$ LG	.11 (.0425)	.101 (.0412)
Gaullist/Domestic Resistance	.453 (.662)	.25 (.823)
Jury $\times$ Gaullist/Domestic Resistance	.351 (.895)	-.145 (1.21)
Gaullist/Communist Resistance	-.122 (.145)	-.0839 (.15)
Jury $\times$ Gaullist/Communist Resistance	.0391 (.214)	.0483 (.235)
Foreign/Domestic Resistance	-.000214 (.000281)	-.00023 (.000347)
Jury $\times$ Foreign/Domestic Resistance	-.000718 (.000305)	-.000708 (.000339)
Communist/Domestic Resistance	-.39 (.149)	-.319 (.163)
Jury $\times$ Communist/Domestic Resistance	.191 (.187)	.155 (.184)
Observations	798	798
Adj R ²	.256	.279
Proof of resistance	Yes	Yes
Individual controls		Yes

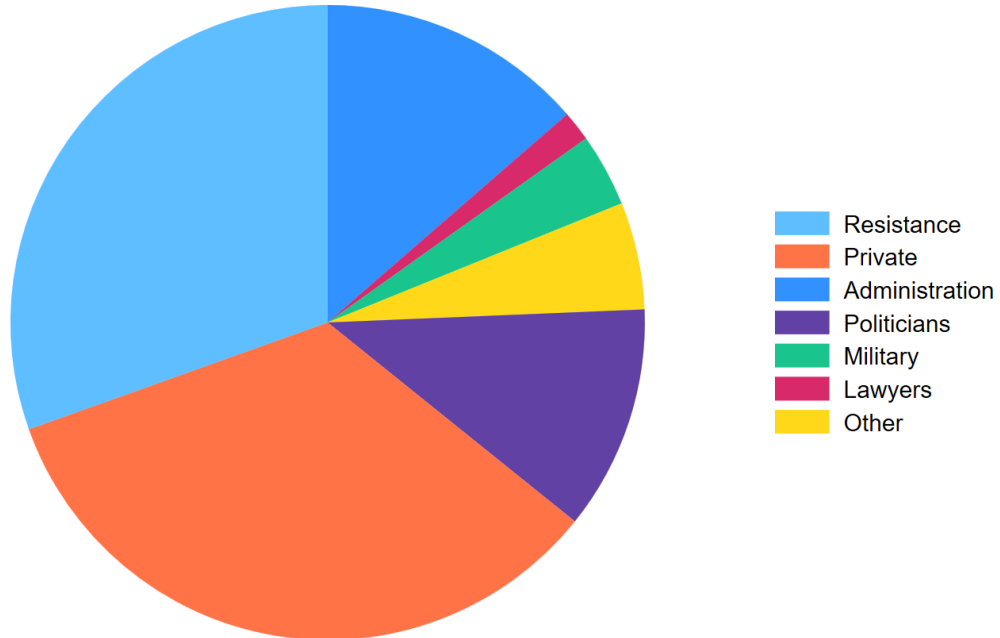
Note: The table reports two specifications of Equation (1) – with and without the baseline control variables and their interaction with the *Jury* dummy variable – augmented with proxies for the composition of the Resistance in the *département* and the interaction between these and the *Jury* dummy variable. The assumption is that the composition of the resistance in a *département* is highly correlated with the composition of the judges on the CDL in that *département*. We measure the composition of the Resistance in a *département* via different ratios of the membership of different resistance groups as defined by certificates of resistance. For example, Gaullist/Communist Resistance is defined as the ratio of the size of the membership of gaullist resistant factions and the membership of communist resistant factions in a *département*. Variables controlling for proof of participation in the civil or military Resistance are included in both specifications. Individual controls include: Age, Jewishness, Journalist, Mayor, Special Role in the Assembly, Arrested by Vichy, Mayor under Vichy, in occupied territory. Standard errors are clustered at the *département* level (in parentheses).

D Additional evidence on the mechanisms

D.1 Presentation - The structure of the dossier data set

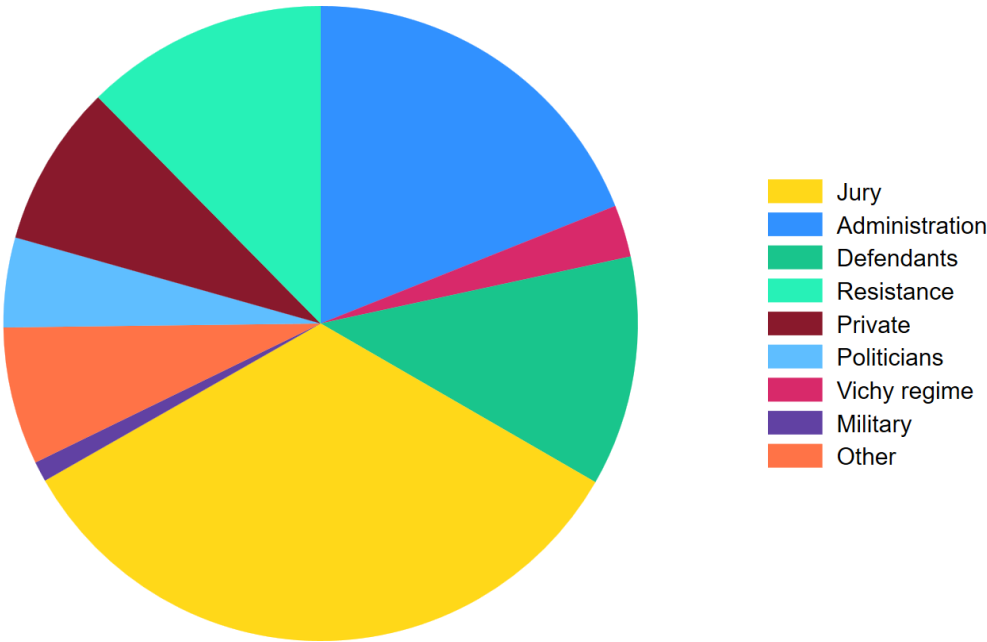
Figure D.1 presents information on the origin of letters of support. These letters represent 19.2% of the documents in a typical dossier. A third of them were private correspondence sent by the defendant’s friends, family, or by individuals in his constituency (“private”) and 30.5% originated from resistance organizations. Figure D.2 presents information on the origin of all documents in the dossiers.

Figure D.1: The origins of letters of support in the dossiers of the defendants



Note: The categories are defined as follows. Administration refers to documents produced by a ministry or a local administration. Vichy Regime refers to documents from the archives of the Vichy regime or from any Vichy-related institutions. Defendant refers to documents produced by the defendant himself. Jury refers to documents produced by the *Jury*. Military refers to documents produced by the French army. Private refers to documents produced by an individual in his/her own name without stating an obvious relation to an organization typically coming from family members or friends of the defendant or from individuals in his constituency. Politicians refers to documents produced by parliamentarians and local politicians. Resistance refers to documents produced by members of resistance networks. Lawyers refers to documents sent by a lawyer using his/her title in the document sent.

Figure D.2: The origins of all the documents in the dossiers of the defendants



Note: The categories are defined as follows. Administration refers to documents produced by a ministry or a local administration. Vichy Regime refers to documents from the archives of the Vichy regime or from Vichy-related institutions. Defendant refers to documents produced by the defendant himself. Jury refers to documents produced by the *Jury*. Military refers to documents produced by the French army. Private refers to documents produced by an individual in his/her own name without stating an obvious relation to an organization, typically coming from family members or friends of the defendant or from individuals in his constituency. Politicians refers to documents produced by parliamentarians and local politicians. Resistance refers to documents produced by members of resistance networks. Lawyers refers to documents sent by a lawyer using his/her title in the document sent.

D.2 Definition and descriptive statistics: The dossier data set

Table D.1: Descriptive statistics - The dossier data set (Part I)

Variable name	Definition	Min	Max	Mean	s.d
General structure					
Nb Doc	Number of documents in the dossier	10	170	40.69	26.38
Nb Pages	Number of pages in the dossier	12	384	55.44	45.26
Nb Doc from Jury	Number of Document produced by the Jury	4	50	13.70	6.16
Nb Archival Docs	Number of Archives	0	50	2.79	6.23
Nb Information requests	Number of information requests sent by the Jury	0	4	1.08	0.95
Nb Letters of support	Number of letters of support	0	90	7.81	12.34
Nb Letters of support - in Favor	In favor of acquitting the defendant	0	90	7.41	11.89
Nb Letters of support - Against	Against acquitting the defendant	0	20	0.22	1.31
Nb Letters of support - Neutral	Neutral	0	10	0.18	0.71
Nb Letters of support - Res	from the Resistance	0	39	2.41	4.28
Nb Letters of support - Mil	from the Military	0	11	0.27	1.01
Nb Letters of support - Administration	from an administration	0	41	1.02	2.91
Nb Letters of support - Others	from other type of organizations	0	70	2.66	6.38
Nb Doc - Military resistance	Nb of Doc mentioning participation in military resistance	0	35	0.77	3.11
Nb Doc - Civilian resistance	Nb of Doc mentioning participation in civilian resistance	0	64	8.55	10.63
Nb Doc - Resistant Press	Nb of Doc mentioning participation in resistant press	0	13	0.59	1.79
Nb Doc - Legal Arguments	Nb of Doc mentioning legal arguments	0	27	5.43	3.95
Nb Doc - Political opinion	Nb of Doc mentioning political opinions of the defendant	0	90	10.02	14.34
Nb Doc - Reelection	Nb of Documents mentioning reelection prospects	0	18	1.07	2.37
Nb Doc - Other topic	Nb of Documents mentioning other topics	0	58	6.32	8.65

Table D1 - Descriptive Statistics - Dossiers data set (Part II)

Direct connections					
1. Nb Doc from defendant	Nb of Docs sent by the Defendants	0	62	4.64	6.05
1.1 Nb of Doc - from Def without legal content	Nb of Doc sent by the Defendant - No legal content	0	19	1.45	2.23
1.2 Nb of Doc - from Def with legal content	Nb of Doc sent by the Defendant - Legal content	0	55	3.19	4.74
1.3 Nb of pages - from Def without legal content	Nb of Pages sent by the Defendant - No legal content	0	53	1.87	3.81
1.4 Nb of pages - from Def with legal content	Nb of Pages sent by the Defendant - Legal content	0	321	9.68	22.05
2. Nb Informal letters from Defendant	Nb of letters with headings "Cher"	0	6	0.17	0.62
2.1 Nb of Doc - Informal letters without legal content	Nb of letters with headings "Cher" - No legal content	0	3	0.06	0.30
2.2 Nb of Doc - Informal letters with legal content	Nb of letters with headings "Cher" - Legal content	0	6	0.11	0.50
Indirect connections					
1. Nb Letters of support from Paris	Nb of letters of support by a Parisian sender	0	38	2.07	4.77
2. Nb of Supporters from Paris	Nb of different supporters from Paris	0	14	0.84	1.81
3. Indirect connections via supporters	Average number of documents sent by different supporters (even outside the case)	0	583.83	33.22	78.38
3. Indirect connections via supporters (excluding letters)	Same as above excluding letters	0	341.67	18.29	46.80
5. Nb Informal documents ("Dear" Letters)	Nb of document with headings "Cher"	0	27	1.46	3.02
5.1 Nb Informal documents to Cassin	Addressed to R. Cassin	0	5	0.19	0.56
5.2 Nb Informal documents to Jury	Addressed to the Jury	0	5	0.22	0.61
5.3 Nb Informal documents not to Jury	Not addressed to the Jury	0	26	1.24	2.84

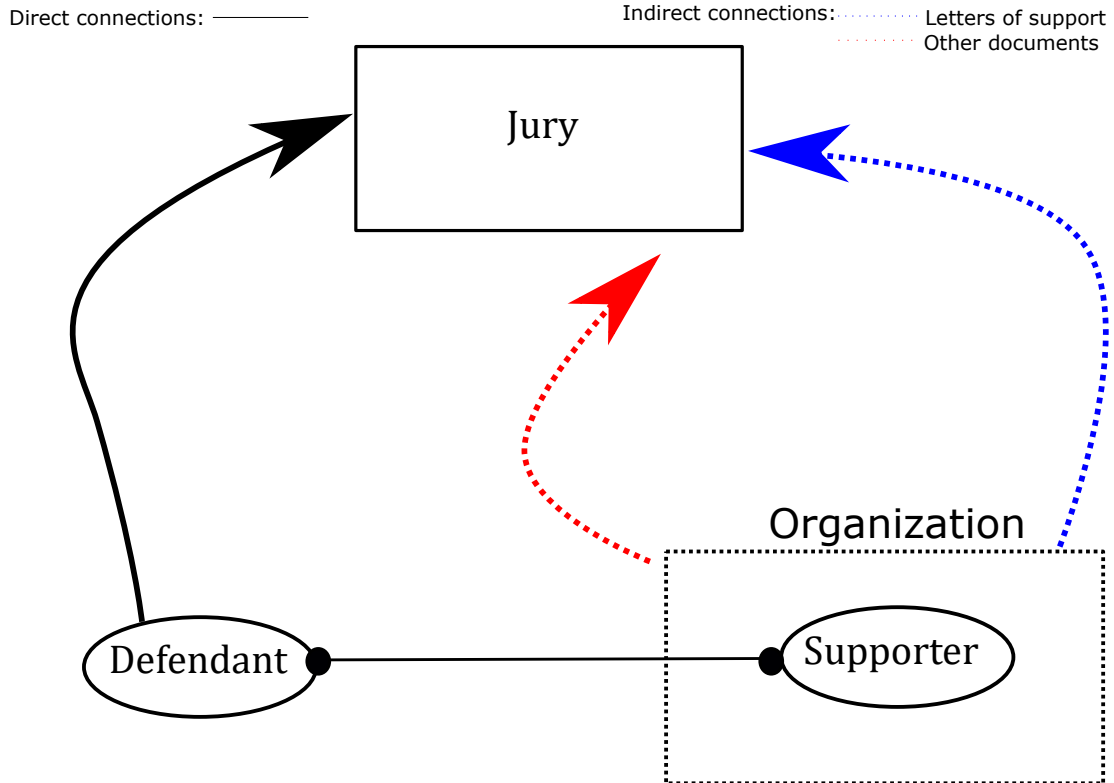
D.3 Measures of indirect connections - Definitions and illustrations

To construct the index of indirect connections, “Indirect connections via supporters”, we single out n organizations across the entire data set with $s = 1, \dots, n$ where n is the total number of organizations in the data set. We weight the share of letters in a defendant’s portfolio from organization s with the total number of documents ($NbDocs_s$) from organization s across all defendants in the entire data set. Formally, for defendant i , the index is defined as

$$\sum_{s=1}^n \left(\frac{NbLetters_{s,i}}{\sum_{s=1}^n NbLetters_{s,i}} \times weight_s \right), \quad (11)$$

where $\frac{NbLetters_{s,i}}{\sum_{s=1}^n NbLetters_{s,i}}$ is the share of letters of support in defendant i ’s portfolio that originated from supporters associated with organization s and $weight_s$ is the weight attached to organization s .³ We set $weight_s = NbDocs_s - Letters_s$ to avoid capturing the efforts of supporters in sending letters but have experimented with alternative weights that add letters of support. It makes little difference. Figure D.3 illustrates the logic behind the index.

Figure D.3: Illustration - Measures of connections



³In cases where a defendant did not get any letters of support, the index takes the value of 0.

D.4 Connections and difference in acquittal rate

D.4.1 The geography of the network

Figure D.4: Distance of the place of birth to Paris
Effect for LG and other defendants

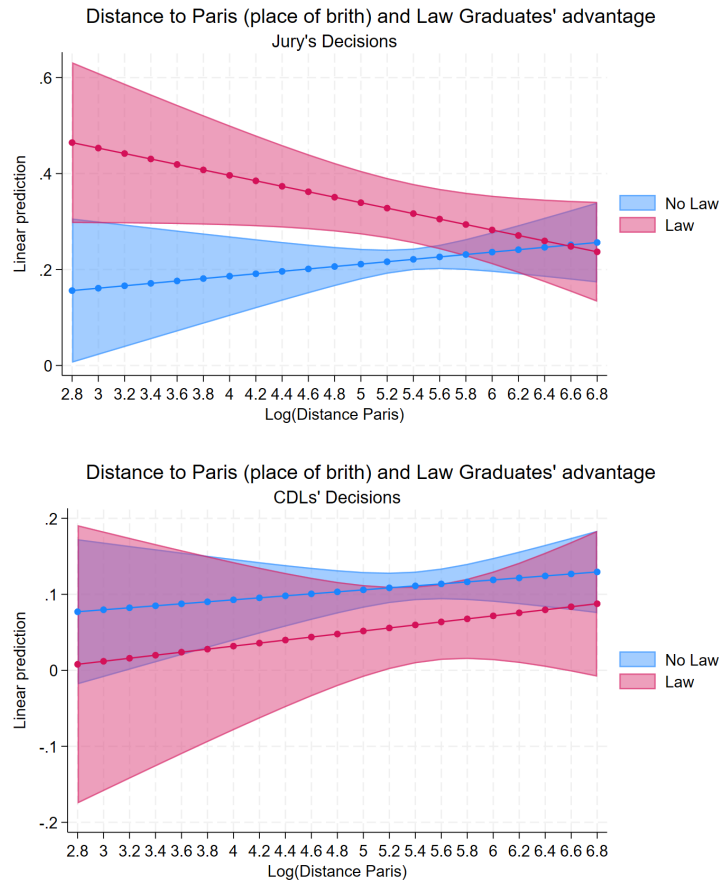


Table D.4: The probability of attending a Parisian Law University for Parisian versus non-Parisian Law graduates

	T-test – Probability(Paris LG=1)		
	Treated	Control	Diff=0
Considered as treated =	Group	Group	(p-value)
Born <100km of Paris	0.94	0.46	0.00
Born <200km of Paris	0.84	0.44	0.00
Born <300km of Paris	0.63	0.46	0.08

Note: This table compares the probability of attending a Parisian Law University by geographical characteristics and where the MP was born. Born <100km of Paris, Born <200km of Paris, Born <300km of Paris are three dummy variables equal to one if the capital city of the birth *département* of the MP is within a 100km, 200km, 300km radius around Paris.

Table D.5: The differential acquittal probability for Parisian versus non-Parisian Law graduates

Panel A: Control group = All those considered non-treated								
Considered as treated =	CDLs			Jury			Diff-in-Diff	
	(D.5.1)	(D.5.2)	(D.5.3)	(D.5.4)	(D.5.5)	(D.5.6)	(D.5.7)	(D.5.8)
	Treated	Control	Diff=0	Treated	Control	Diff=0	$\Delta\Delta$	p-value
	Group	Group	(p-value)	Group	Group	(p-value)		
All Law graduates	0.08	0.11	0.45	0.31	0.23	0.097	0.11	0.03
Law graduates (Parisian U)	0.07	0.11	0.37	0.36	0.23	0.04	0.16	0.01
Other Law graduates	0.10	0.10	0.96	0.25	0.25	0.94	0.01	0.91
Panel B: Control group = All non-Law graduate								
Considered as treated =	CDLs			Jury			Diff-in-Diff	
	Treated	Control	Diff=0	Treated	Control	Diff=0	$\Delta\Delta$	p-value
	Group	Group	(p-value)	Group	Group	(p-value)		
Law graduates (Parisian U)	0.07	0.11	0.36	0.36	0.23	0.04	0.17	0.01
Other Law graduates	0.10	0.11	0.84	0.26	0.23	0.68	0.04	0.58

Note: The estimates in Panel A make use of the whole sample, i.e., they use all non-treated defendants as controls. The estimates in Panel B compares subgroups of Law graduates to non-Law graduates, i.e., they exclude non-treated Law graduates from the control group. Column D.5.1 reports the average acquittal rate of the treated group (defined in the left column) in front of the CDLs whereas column D.5.2 reports the average acquittal rate of the control group (i.e., all individuals not in the treated group) in front of the CDLs. Column D.5.3 displays the difference between these two means. Column D.5.4 presents the average acquittal rate of the treated group in front of the *Jury* whereas column D.5.5 presents the average acquittal rate of the control group (i.e., all individuals not in the treated group) in front of the *Jury*. Column D.5.6 displays the difference between columns D.5.4 and D.5.5. Column D.5.7 reports the difference-in-differences estimate from Equation (1) (without any control) of the difference in acquittal rate between each of the two subgroups of Law graduates defined in the left column and other defendants. Standard errors are clustered at the département level.

D.4.2 Connections - Parisian and non-Parisian Law graduates

Table D.6: Documents in the court dossiers: Parisian versus non-Parisian Law graduates

Control group = All considered as non-treated						
	Mean			Mean		
	(D.6.1)	(D.6.2)	(D.6.3)	(D.6.4)	(D.6.5)	(D.6.6)
	LG	Control	Diff=0	LG	Control	Diff=0
	(Paris U)			(not Paris U)		
Difference in:	Group	Group	(p-value)	Group	Group	(p-value)
Nb Letters of support from Paris	3.63	1.80	0.01	1.69	2.12	0.54
Nb of Supporters from Paris	1.59	0.71	0.00	0.80	0.85	0.86
Indirect connections via supporters	57.01	29.09	0.01	36.31	32.77	0.76
Indirect connections via supporters (excl. letters)	32.73	15.78	0.01	19.48	18.11	0.85
Nb Informal documents ("Dear" Letters)	2.78	1.24	0.00	1.08	1.52	0.33
Nb Informal documents to Cassin	0.19	0.19	0.98	0.12	0.20	0.34
Nb Informal documents to Jury	0.27	0.21	0.49	0.16	0.23	0.42
Nb Informal documents not to Jury	2.51	1.02	0.00	0.92	1.29	0.39

Control group = All non law graduates						
	Mean			Mean		
	LG	Control	Diff=0	LG	Control	Diff=0
	(Paris U)			(not Paris U)		
Difference in:	Group	Group	(p-value)	Group	Group	(p-value)
Nb Letters of support from Paris	3.63	1.82	0.01	1.69	1.82	0.85
Nb of Supporters from Paris	1.59	0.70	0.00	0.80	0.70	0.65
Indirect connections via supporters	57.01	27.82	0.01	36.31	27.82	0.46
Indirect connections via supporters (excl. letters)	32.73	15.13	0.01	19.48	15.13	0.53
Nb Informal documents ("Dear" Letters)	2.78	1.26	0.00	1.08	1.26	0.63
Nb Informal documents to Cassin	0.19	0.20	0.86	0.12	0.20	0.35
Nb Informal documents to Jury	0.27	0.22	0.58	0.16	0.22	0.49
Nb Informal documents not to Jury	2.51	1.04	0.00	0.92	1.04	0.73

Note: The upper panel provides estimates on the whole sample. The bottom panel provides estimates of one subgroup of Law graduates compared to another. Column D.6.1 presents the averages of the various measures of document content for Law graduates from a Parisian University (Paris U); Column D.6.4 displays averages for Law graduates from other universities (not Paris U). Columns D.6.2 and D.6.5 show the averages for control groups. Columns D.6.3 and D.6.6 show the difference between averages of the treated and the control group. Standard errors are clustered at the département level.

D.4.3 Difference in the treatment of information between CDLs and the Jury - A Law graduate effect?

Table D.7: Did the CDLs treat information related to the actions and events during the war differently for Law graduate?

Dep Variable	(D.7 .1)	(D.7.2)	(D.7.3)	(D.7.4)	(D.7.5)
Sample	<i>Acquit_i</i> Decisions CDL	<i>Acquit_i</i> Decisions CDL	<i>Acquit_i</i> Decisions CDL	<i>Acquit_i</i> Decisions CDL	<i>Acquit_i</i> Decisions CDL
LG × Civil Res	.131 (.0723)				.133 (.0639)
LG × Military Res		.0155 (.118)			.0148 (.116)
LG × Arrested Vichy			.39 (.244)		.374 (.252)
LG × Mayor Vichy				.00883 (.0797)	.0307 (.0716)
Observations	395	395	395	395	395
Adj R ²	.0595	.11	.0566	.0657	.138
Individual controls	Yes	Yes	Yes	Yes	Yes
Dep FE + Interaction	Yes	Yes	Yes	Yes	Yes

Note: Linear probability models estimated on the sample of 395 decisions made by the CDLs. The dependent variable is the acquittal decision by the CDL for each defendant. The purpose is to assess if the CDLs reacted to information contained in the dossiers of the *Jury* in a different way for Law graduates (LG) than for other defendants. We interact the Law graduate dummy variable with four variables relating to war-related events recorded in the dossiers: participation in the civil or military resistance, arrest by the Vichy regime (Vichy), and served as mayor under the Vichy regime. Individual controls include: Age, Jewishness, Journalist, Mayor, Special Role in the Assembly, MP of an occupied territory. Standard errors are clustered at the département level.

Table D.8: Difference in the treatment of information between CDLs and the Jury

	(D.8.1)	(D.8.2)	(D.8.3)	(D.8.4)	(D.8.5)
Dep Variable			<i>Acquit_{i,c}</i>		
Sample	All	All	All	All	All
LG $\times$ Jury	.143 (.0594)	.155 (.06)	.132 (.0532)	.147 (.0594)	.162 (.0543)
Civil Res $\times$ Jury	.252 (.0421)				.201 (.0444)
LG $\times$ Civil Res $\times$ Jury	-.0401 (.0931)				-.0389 (.108)
Mili Res $\times$ Jury		.269 (.0778)			.226 (.0775)
LG $\times$ Mili Res $\times$ Jury		-.00753 (.166)			.00326 (.168)
Arrested Vichy $\times$ Jury			.239 (.11)		.238 (.108)
LG $\times$ Arrested Vichy $\times$ Jury			.0715 (.308)		.0782 (.278)
Mayor Vichy $\times$ Jury				.0273 (.0584)	.0415 (.0545)
LG $\times$ Mayor Vichy $\times$ Jury				-.0436 (.119)	-.0148 (.114)
Observations	790	790	790	790	790
Adj R ²	.178	.237	.12	.112	.297
Controls	Yes	Yes	Yes	Yes	Yes

Note: This table estimates a version of equation (1) augmented with triple interactions of the Jury dummy variable (Jury), the Law graduate dummy variable (LG), and the different measures of participation (or non-participation) in the resistance. The sample include all decisions (i.e., both by the CDLs and the *Jury*). All regressions include LG, Jury and LG $\times$ Jury (not reported). The purpose of the regressions is to investigate the specific reaction of the *Jury* to this type of information for Law graduates relative to other defendants. Individual controls include: Age, Jewishness, Journalist, Mayor, Special Role in the Assembly, MP of an occupied territory. Each individual control is also interacted with the Jury dummy variable. Standard errors are clustered at the département level (in parentheses)

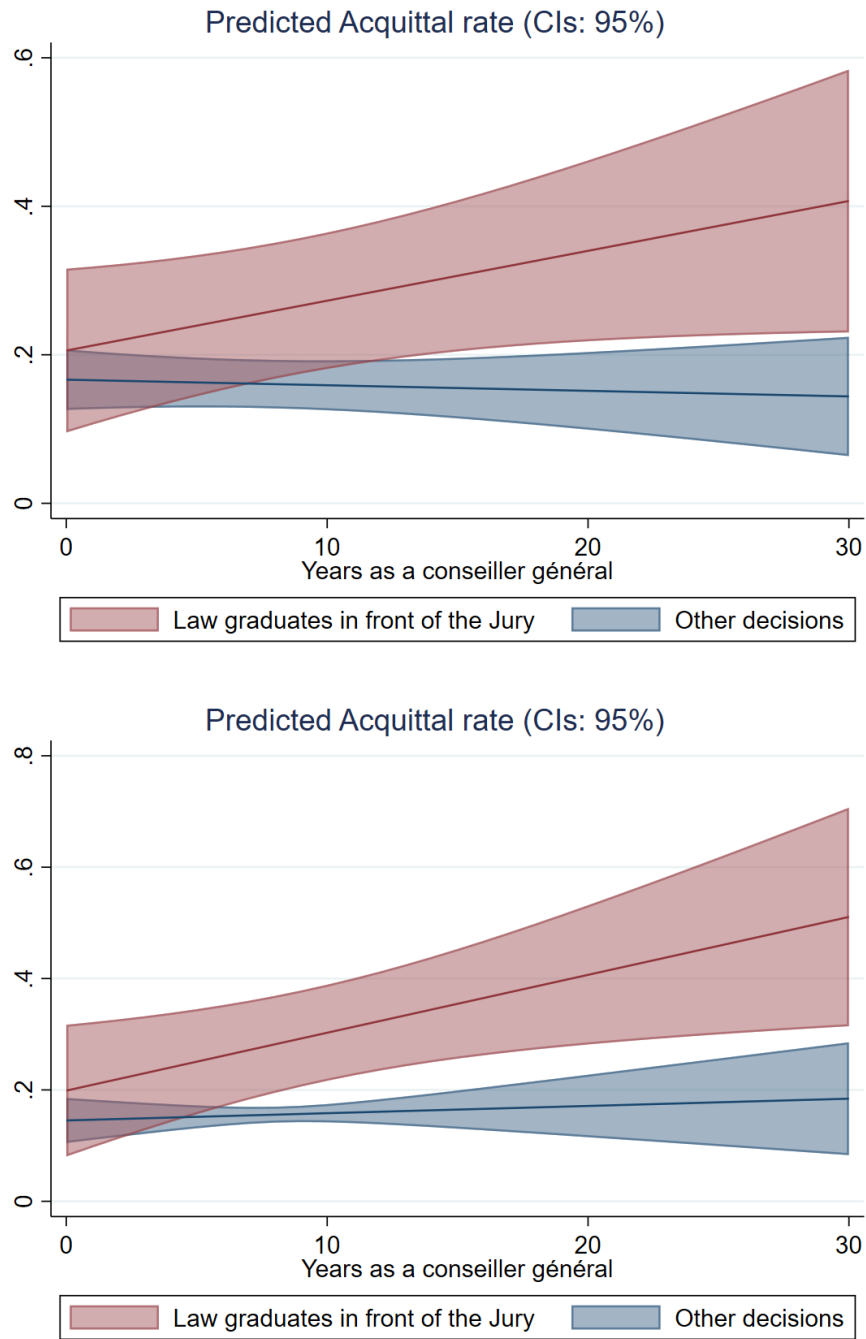
D.4.4 The higher acquittal of Law graduates is not explained by legal experience

Table D.9: Legal experience and acquittal rate

Dependent variable	(D.9.1) Diff-in-diff	(D.9.2) Diff-in-diff
	$Acquit_{i,c}$	$Acquit_{i,c}$
Jury	.155 (.0286)	-.17 (.169)
LG	-.0174 (.0383)	-.0201 (.0519)
LG $\times$ CG	-.00115 (.00309)	-.00265 (.00413)
LG $\times$ Jury	.0392 (.0598)	.0536 (.0662)
LG $\times$ Jury $\times$ CG	.00747 (.00385)	.00909 (.0047)
Observations	798	798
Adj R ²	.0399	.302
Full controls	No	Yes

Note: The table reports specifications of equation (1) augmented with a triple interaction between the length of the political career of a defendant before the war and LG $\times$ Jury. The hypothesis is that a long political career crowds out legal skills. Column D.9.2 adds the full set of individual control variables. Individual controls include: Age, Jewishness, Journalist, Mayor, Special Role in the Assembly, MP of an occupied territory. Each individual control is also interacted with the Jury dummy variable. Standard errors are clustered at the département level. (in parentheses)

Figure D.5: Interaction plots: the probability of acquittal as a function of years of service as a local councilor



Note: The two figures exhibit interaction graphs based on the estimates in Table D.9, columns D.9.1 (upper graph) and D.9.2 (lower graph). The red line is for Law graduates in front of the *Jury*; the blue line is for other decisions (related to non-Law graduate defendants in front of the *Jury* and CDLs).

D.5 Correlation Matrix - Measures of indirect connections

Table D.10: Correlation matrix - Measures of indirect connections

	1. Nb Letters of support from Paris	2. Nb of Supporters from Paris	3. Indirect connections via supporters	4. Indirect connections via supporters (excl. letters)	5. Nb Informal documents ("Dear" Letters)	5. Nb Informal documents ("Dear" Letters)
1. Nb Letters of support from Paris	\					
2. Nb of Supporters from Paris	0.82	\				
3. Indirect connections via supporters	0.55	0.34	\			
4. Indirect connections via supporters (excl. letters)	0.52	0.32	0.99	\		
5. Nb Informal documents ("Dear" Letters)	0.55	0.52	0.37	0.35	\	
5. Nb Informal documents not Jury ("Dear" Letters)	0.54	0.52	0.37	0.35	0.98	\

E Elite persistence after the purges

E.1 Estimates of persistence

For the difference in acquittal rates between Law graduates and other defendants to have contributed to elite persistence, it must have helped the future careers of the politicians who benefited from it. Decisions of the *Jury* would have been inconsequential if acquitted Law graduates never ran for election or if voters did not elect them. In this section, we show that the decisions of the *Jury* were consequential for the post-war careers of defendants and that Law graduates were more likely than other defendants to pursue a career in politics after being cleared by the *Jury*. Table E.1 reports a series of regressions in which measures of each defendant's post-war political career (such as the number of times he ran in municipal and legislative elections, whether he was mayor, held a seat in parliament or had ministerial responsibilities) are explained by the court decisions (Panel A) or by being a Law graduate (Panel B). We find very similar results when looking at the intensive margin of the effect in Table E.2.

Table E.1: Elite persistence - the *Jury*'s decisions and Law graduates

Panel A - Independent variable: Acquitted _{Jury}					
	(E.1.A.1) Nb mayor election	(E.1.A.2) Nb deputy elections	(E.1.A.3) Mayor=1	(E.1.A.4) Parliament=1	(E.1.A.5) Minister=1
Acquitted _{Jury}	.189 (.071)	.253 (.0705)	.116 (.0515)	.225 (.0526)	.0701 (.0309)
Acquitted _{CDL}	.0431 (.097)	.0735 (.1)	.0115 (.064)	.0553 (.0771)	.00894 (.0399)
Observations	397	399	399	399	399
Adjusted R ²	.172	.156	.166	.157	.0671
<u>Control variables</u>					
Individual	YES	YES	YES	YES	YES

Panel B - Independent variable: Law graduate					
	(E.1.B.1) Nb mayor elections	(E.1.B.2) Nb deputy elections	(E.1.B.3) Mayor=1	(E.1.B.4) Parliament=1	(E.1.B.5) Minister=1
LG	-.0352 (.0438)	.105 (.0476)	-.0107 (.0348)	.0761 (.0348)	.0318 (.0181)
Acquitted _{CDL}	.121 (.0977)	.188 (.0961)	.0597 (.0614)	.156 (.0829)	.0408 (.0374)
Observations	397	399	399	399	399
Adjusted R ²	.153	.125	.151	.11	.0436
<u>Control variables</u>					
Individual	YES	YES	YES	YES	YES

Note: Level of observation: Defendant. In Panel A, Acquitted_{Jury} is a dummy variable equal to 1 if the defendant was acquitted by the *Jury*. In Panel B, Law graduate is a dummy variable equal to one if the defendant was a Law graduate. Acquitted_{CDL} is a dummy variable equal to one if the defendant was acquitted by a CDL. Dependent variables: Nb mayor elections is the number of times a defendant was a candidate in a municipal election after the war. Nb deputy elections is the number of times a defendant was a candidate in a legislative election after the war. Mayor=1 is a dummy variable equal to one if a defendant was elected mayor after the war and zero otherwise. Parliament=1 is a dummy variable equal to one if a defendant was a deputy or a senator after the war and zero otherwise. Minister=1 is a dummy variable equal to one if the defendant was a minister after the war and zero otherwise. Individual controls include: Age and Jewishness); Journalist; Political mandates (Mayor, Special Role in the Assembly, parliamentarian of an occupied territory); Resistance and collaboration WWII (Civilian Resistance, Military Resistance, Arrested by the Vichy regime, Mayor under Vichy). Standard errors are clustered at the département level.

Table E.2: Persistence and the *Jury's* decision - the intensive margin

Panel A - Independent variable: Acquitted by the Jury			
	(E.2.A.1) Years as Mayor	(E.2.A.2) Years in Parliament	(E.2.A.3) Years as Minister
Acquitted _{Jury}	.271 (.135)	.366 (.104)	.0695 (.0336)
Acquitted _{CDL}	-.0373 (.163)	.117 (.149)	.0344 (.0611)
Observations	399	399	399
Adjusted R ²	.155	.137	.0508
<u>Control variables</u>			
Individual	YES	YES	YES

Panel B - Independent variable: Law graduates			
	(E.2.B.1) Years as Mayor	(E.2.B.2) Years in Parliament	(E.2.B.3) Years as Minister
LG	-.016 (.0847)	.181 (.071)	.0341 (.0247)
Acquitted _{CDL}	.0763 (.159)	.284 (.155)	.0662 (.0607)
Observations	399	399	399
Adjusted R ²	.142	.115	.0378
<u>Control variables</u>			
Individual	YES	YES	YES

Note: Level of observation: Defendant. In Panel A, the main independent variable is a dummy variable equal to 1 if the defendant has been acquitted by the Jury. In Panel B, the main independent variable is a dummy variable equal to one if the defendant was a law graduate. Dependent variables: Years as Mayor is the number of years the defendant has spent as mayor after WWII (Log-transformed). Years in Parliament is the number of years the defendant has spent either as a deputy or as a senator after WWII (Log-transformed). Years as Minister is the number of years the defendant has spent as a minister after WWII (Log-transformed). Individual controls include: Age and Jewishness; Journalist; Political mandates (Mayor, Special Role in the Assembly, parliamentarian of an occupied territory); Resistance and collaboration WWII (Civilian Resistance, Military resistance, Arrested by the Vichy regime, Mayor under Vichy). Standard errors are clustered at the département level.

E.2 Willingness to continue in politics and the differential acquittal rate of Law graduates in front of the *Jury*.

We proxy the intentions of defendants in two ways. First, we use the information in their dossiers related to participation in the 1945 municipal elections and divide the defendants into two groups: those who ran for election and therefore intended to pursue a political career, and those who did not. Table E.3, columns E.3.1 and E.3.2 report separate estimates of equation (1) for the two groups. We observe that the difference in acquittal rate between Law graduates and other defendants is only significant, at the ten-percent level, for the group of defendants who ran in the 1945 municipal election. Moreover, Column E.3.3 reports a specification estimated on the full sample in which equation (1) is augmented with a triple interaction between the Law graduate dummy, the *Jury* dummy, and a dummy coding whether the defendant ran in the election. The coefficient on the triple interaction term is positive and significant at the one-percent level. Its magnitude indicates that the acquittal rate before the *Jury* was 40 percentage points higher for Law graduates who ran for election than for those who did not.

Second, as an alternative proxy for a defendant's intention to seek public office, we leverage a discontinuity caused by a change in the remit of the *Jury*. Until September 1945, the *Jury* was in charge of judging two types of cases: cases of electoral litigation brought by departmental prefects and cases brought by the defendants themselves. These cases were mainly about eligibility to run for election. An order of 13 September 1945 expanded the remit of the *Jury* to include the cases of all parliamentarians who had voted in favor of the enabling act or had collaborated with the Vichy regime. Many of whom did not intend to run in an election. We know from the dossiers of the defendants when a case was considered by the *Jury* and can, therefore, distinguish cases considered before and after the change in the remit and create a pre- and a post-reform dummy. In this way, we can use the discontinuity to test if the difference in acquittal rate between Law graduates and other defendants before the *Jury* was larger for defendants who wanted to continue their political career (as revealed by an early case related to eligibility for election) than for other defendants. Specifically, we augment equation (1) with interaction terms between the pre- and post-reform dummies and $LG_i \times Jury_c$, respectively, to allow the coefficient β_1 to differ depending on when the case was heard. In Appendix Table E.3, columns E.3.4 to E.3.6 present the results. The result in column E.3.4 shows that the difference in acquittal rate between Law graduates and other defendants before the *Jury* was around 29 percentage points larger for the defendants tried before the reform when the *Jury* focused on electoral litigation than for those tried after the

expansion of its remit.⁴ A Wald-test shows that this difference is statistically significant. The specifications in the other two columns are augmented with time polynomials and their interactions with $LG_i \times Jury_c$ and the reform dummy variables and show that the effect is not driven by time trends in the sentences of the *Jury*.⁵ The difference in acquittal rate between Law graduates and other defendants before the *Jury* therefore facilitated elite persistence, as it materialized specifically when a Law graduate intended to continue his political career.

Table E.3: Higher acquittal of Law graduates when it matters: Electoral litigations

Dep variable	(E.3.1)	(E.3.2)	(E.3.3)	(E.3.4)	(E.3.5)	(E.3.6)
Sample	$Acquit_{i,c}$ Ran for elections	$Acquit_{i,c}$ Did not run	$Acquit_{i,c}$ All	$Acquit_{i,c}$ All	$Acquit_{i,c}$ All	$Acquit_{i,c}$ All
LG	-.221 (.448)	.000193 (.0405)	-.0111 (.04)			
$Jury \times LG$	.888 (.146)	.0751 (.052)	.0733 (.0526)			
$Jury \times LG \times$ Ran for elections			.558 (.195)			
$Jury \times LG \times$ Pre-reform				.381 (.136)	.332 (.122)	.369 (.128)
$Jury \times LG \times$ Post-reform				.0927 (.0517)	.114 (.0516)	.133 (.0496)
Observations	66	658	790	790	784	784
Adjusted R ²	.289	.311	.303	.308	.349	.359
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dep FE + Interaction	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial Date					3	3
Pre-reform dummy						Yes
Wald Test				0.06	0.12	0.10

Note: Column E.3.1 estimates equation (1) on the subset of defendants who ran for the first post-WWII mayoral elections. Column E.3.2 estimates equation (1) on the subset of defendants who ran for the first post-WWII mayoral elections. Column E.3.3 investigates how the magnitude of the coefficient β_1 varied with their participation in the first post-WWII elections by adding a triple interaction term ($Jury \times LG \times$ Ran for elections) and controlling for the interaction $Jury \times$ Ran for elections. It controls for the variables not interacted in the estimation ($Jury$, LG , and Ran for elections). Columns E.3.4 to E.3.6 estimate equation (6) in a manner akin to a RDD estimates using the cutoff of the September 13rd as a discontinuity. It therefore shows how the reform affected the bias of the *Jury* towards Law graduates after adding individual controls, a time-polynomial of order 3, and a pre-reform dummy variable. This test assesses whether the reform of the remit of the *Jury* from mainly electoral litigations to investigations of all cases impacted the time trend of our estimates of the difference in acquittal rates between Law graduates and other defendants. Individual controls include: Age, Jewishness, Journalist, Mayor, Special Role in the Assembly, Civilian Resistance, Military resistance, Arrested by by the Vichy regime, Mayor under Vichy, parliamentarian of an occupied territory. Each invidual control is interacted with the *Jury* dummy variable. Standard errors are clustered at the département level (in parentheses).

⁴To determine the difference between the two groups we compare the coefficients attached to Law graduates tried before the reform and those tried after the reform, hence: 33.8-4.6.

⁵Part of this difference in sentencing patterns over time might be captured by the comparison with the decision of CDLs (also varying over time). Adding time polynomials however allows to take time varying dynamics of the *Jury* into account as previous research has shown that time-dependence exists in sentencing (Bindler and Hjalmarsson, 2018).

E.3 Differences between acquitted Law graduates and other acquitted defendants

Table E.4: Differences between acquitted LG and other acquitted politicians

	<u>All cleared defendants</u>			<u>Not cleared by CDL</u>		
	Cleared	Cleared	Δ	Cleared	Cleared	Δ
	LG	Other		LG	Other	
WWI War Medal	0.50	0.29	-0.21	0.52	0.27	-0.25
Mandate (years)	14.24	9.89	-4.34	13.67	9.71	-3.96
Conseiller général (years)	11.88	6.86	-5.02	11.26	5.66	-5.60
Deputé (years)	10.38	7.76	-2.62	9.44	8.15	-1.30
Senator (years)	3.85	2.14	-1.72	4.22	1.56	-2.66

Note: WWI War Medal is a dummy variable equal to one if the defendant was awarded a war medal for his action during Word War II. Mandate (years) is the number a defendant was in parliament (either as a Député or as a Senator. Conseiller général (years) is the number of years a defendant has been a conseiller général (member of départemental assemblies). Deputé (years) is the number of years a defendant acted as Deputé and Senator (years) is the number of years a defendant acted as Senator.

F Descriptive statistics - Different data sets

F.1 Descriptive statistics related to defendant characteristics

Table F.1: Descriptive statistics - Defendant characteristics (Part I)

Variable	Mean	s.d	Min	Max
Law graduates	0.28	0.45	0	1
Dependent variable				
Acquittal in front of Jury	0.25	0.43	0	1
Acquittal in front of CDL	0.10	0.30	0	1
Politics and political mandates				
Senator	0.35	0.48	0	1
Rightwing	0.51	0.50	0	1
Center	0.21	0.41	0	1
Mayor	0.47	0.50	0	1
Pres/Vice-Pres or Sec Assembly	0.08	0.27	0	1
MP elected in Paris	0.05	0.22	0	1
MP of an occupied department	0.52	0.50	0	1
Dynastic politicians	0.16	0.37	0	1
War experience				
WWI Veteran	0.51	0.50	0	1
WWII fighter	0.06	0.24	0	1
Networks, clubs and religion				
Free Mason	0.04	0.19	0	1
Jewish MP	0.02	0.12	0	1
Labour unions	0.08	0.27	0	1
Agricultural organizations	0.10	0.29	0	1
Légion d'Honneur	0.38	0.49	0	1
War Medal	0.38	0.49	0	1
Veterans club	0.05	0.21	0	1

Table F.1: Descriptive statistics - Defendant characteristics (Part II)

Variable	Mean	s.d	Min	Max
Occupation				
Civil Servant	0.06	0.24	0	1
Workers	0.07	0.26	0	1
Journalist	0.12	0.33	0	1
Informational cues				
Excluded by his party	0.27	0.44	0	1
Signed Bergery motion	0.13	0.33	0	1
Mayor under Vichy	0.27	0.45	0	1
Arrested by Vichy	0.06	0.23	0	1
Militarian resistance	0.20	0.40	0	1
Civilian resistance	0.58	0.49	0	1
Continuous variables				
Age	60.25	10.43	34	84
National Mandate	11.19	8.07	1	38
Conseiller Général	8.63	10.67	0	42
Study Years	3.42	3.14	0	8

F.2 Descriptive statistics related to dossier data set

Table F.2: Descriptive statistics - the dossier data set

Documents	Total	Min per dossier	Max per dossier	Mean per Dossier	s.d
All type of document	17589	1	170	40.62	26.9
Produced by					
Jury	5882	0	50	13.58	6.14
Defendant	2061	0	62	4.76	6.27
Administrations	3335	0	48	7.70	6.00
Politicians	804	0	26	1.86	3.26
Resistant Organizations	2176	0	45	5.03	5.6
Journal Officiel	441	0	3	1.02	0.38
Press	251	0	24	0.58	1.96
Private	1452	0	70	3.35	7.28
Military	181	0	12	0.42	1.34
Intelligence Services	162	0	9	0.37	0.96
Lawyer	115	0	10	0.27	1.03
Court	131	0	6	0.30	0.96
Labor Unions	38	0	4	0.09	0.39
Veterans Association	41	0	7	0.09	0.56
Collaborationist administration	466	0	19	1.08	2.39
Others	54	0	5	0.12	0.52
Type of documents					
Letters of support	3385	0	90	7.82	12.23

References

- Alesina, A. and E. La Ferrara (2014). A test of racial bias in capital sentencing. *American Economic Review* 104(11), 3397–3433. 2
- Bindler, A. and R. Hjalmarsson (2018). Path dependency in jury decision making. *Journal of the European Economic Association* 17(6), 1971–2017.
- Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica* 50(6), 1431–1451. 4
- Grossman, G. M. and E. Helpman (2002). *Special Interest Politics*. MIT Press. 5