Did the Independence of Judges Reduce Legal Development in England, 1600–1800?

Peter Murrell *University of Maryland*

Abstract

Conventional wisdom confers iconic status on the clause of England's Act of Settlement (1701) mandating secure tenure for judges. This paper uses new databases of judges' biographies and citations to estimate how the move to secure tenure affected the number of citations to judges' decisions, a measure of the quality of decisions. Several strategies facilitate identification of the effect of secure tenure. A court-year panel permits use of a difference-in-differences framework. Controls capture judges' human capital and amount of litigation. Historical evidence, tests of sensitivity to omitted-variable bias, and instrumental variables estimates support the findings on the effects of tenure arrangements derived from ordinary least squares estimates. Secure tenure had a strong deleterious effect on associate judges' decisions and a smaller positive effect on chief judges' decisions. The effect of all judges having secure tenure is negative, large, and statistically significant. The act had an effect opposite of that universally assumed.

> According to the plan of the convention, all judges who may be appointed by the United States are to hold their offices *during good behavior*. . . . The standard of good behavior for the continuance in office of the judicial magistracy, is certainly one of the most valuable of the modern improvements in the practice of government. (Hamilton [1788] 1981, pp. 226–27)

> In one sense you could say that Judge [Roy] Bean was independent. He did whatever he liked. (O'Connor 2009, p. 47)

1. Introduction

When referring to one of the most valuable modern improvements in government, no doubt Alexander Hamilton had in mind the clause of the English Act

This is a revised version of Murrell (2017), which was first presented at the 2017 Society for Institutional and Organizational Economics conference. Participants at that conference provided helpful suggestions. I thank Daniel Berkowitz, Dan Bogart, Ken Coriale, Felipe González, Peter Grajzl, Melissa Kearney, and Phil Keefer for helpful comments and suggestions. Nona Karalashvili, Erica Ryan, and Martin Schmidt provided both research assistance and valuable comments.

[*Journal of Law and Economics,* vol. 64 (May 2021)]
 0022-2186/2021/6403-0018$10.00

of Settlement of 1701 that states, "[J]udges commissions be made *Quamdiu se bene gesserint,*" that is, during good behavior (Berger 1970, p. 1500 n. 127). This mandated secure tenure for judges. In the 17th century, English judges could be appointed *durante bene placito regis,* that is, during the king's pleasure, which meant that often they could be removed at the whim of the monarch: they had insecure tenure.[1] The change in tenure arrangements mandated by the act is universally regarded as marking a watershed in legal history, a defining event in the establishment of judicial independence.

To date, Hamilton's hypothesis on the beneficial effect of the secure-tenure clause of the Act of Settlement has never been subjected to empirical tests.[2] This paper does so, using the number of citations to court decisions as a measure of the quality of those decisions. The results are remarkably at odds with almost universal assumptions about English history. Secure tenure has a substantively large, statistically significant, deleterious effect on the citations to decisions of associate judges. The estimated effect of chief-judge secure tenure on later citations is positive but always substantively small and usually statistically insignificant. The associate-judge effect dwarfs that for the chief judges, which means that the effect of all judges having secure tenure on citations is negative, large, and statistically significant.

Even if one views the results of this paper in the most conservative fashion, they present a picture that is starkly at odds with assumptions that today go unquestioned. The results reject Hamilton's hypothesis, which is a fundamental element of the conventional wisdom on 17th- and 18th-century English history, an era that provides a central paradigm for celebrated works discussing modernization and development (for example, North and Weingast 1989; Olson 1993; Acemoglu and Robinson 2012). One might even conclude that secure tenure slowed the development of law. Citations are a measure of how much the individual decisions of a court have contributed to increases in the stock of precedents that constitute a significant part of the common law. The secure-tenure clause of the Act of Settlement thus reduced legal development in England by as much as 20 percent according to some rough calculations presented in Section 6. The remainder of this section summarizes the steps by which these conclusions are reached.

In the standard telling, the act gave judges secure tenure, which made them more independent of the monarch.[3] Then judicial performance improved, which

[1] *Quamdiu se bene gesserint* is literally translated as "as long as he should conduct himself well" but is usually rendered as "during good behavior." *Durante bene placito* is literally translated as "during the pleasure [of the grantor of the office]" but is usually rendered in this context as "during the king's pleasure." The official name of the act is An Act for the Further Limitation of the Crown and Better Securing the Rights and Liberties of the Subject. For the wording of the act, see Raithby (1820, pp. 636–38).

[2] Klerman and Mahoney (2005) show that financial returns varied with subjective probabilities of the passage of the relevant clause of the act. This is not a test of the effect of the independence clause but rather is about the prevailing expectations of its effect, that is, the conventional wisdom of that era.

[3] Secure tenure conditioned only on good behavior is a core ingredient of formal independence and an important determinant of actual independence (Feld and Voigt 2003).

established a core element of the legal basis for a modern economy.[4] This story has three links: the direct effect of the act on formal judicial institutions, the subsequent effect of the institutions on judicial behavior, and the resultant effect of that behavior on English legal development. This paper examines the first two links, which provides evidence pertinent to the third.

The standard story concerning the first link omits important nuances. The use of secure tenure was frequent before the Act of Settlement. Because the act did not apply to all judges, insecure tenure continued after it was passed.[5] Both forms of judicial tenure coexisted for the whole of the 17th and 18th centuries, sometimes even simultaneously for different judges on the same courts. Section 2 provides an overview of the facts about the institutions relevant to judicial tenure over the period 1600–1800, integrating evidence that has hitherto been scattered and often inaccurately portrayed in the literature.

With the more nuanced picture revealing that there was significant variation in the terms of the appointments of judges, it becomes clear that there is an opportunity for empirical work on the second link, the one from legal institutions to judicial behavior. Because the two types of judicial appointments existed side by side and varied in relative importance both within time periods and within courts, it is possible to identify the effect of judicial tenure arrangements on judicial performance. The measure of judicial performance used here is the number of citations to the reports of cases in which judges participated, a standard measure of the quality of judicial decisions.

The empirical investigation uses two new databases, described in Online Appendix OB. Each contains data on English history that were previously dispersed in several sources in nondigital form, not readily available for empirical work. The first provides biographical information about all judges serving on England's highest courts in the 17th and 18th centuries. The crucial explanatory variables are the provisions of their appointments. But this database also provides useful controls such as proxies for human capital. The second data set documents all citations to reports of previous cases that appear in *The English Reports* (Renton

[4] On the period in question, see, for example, North and Weingast (1989, pp. 816–20) and Olson (1993, p. 574). Despite some cautionary lessons from the recent literature (for example, Hanssen 2004; Choi, Gulati, and Posner 2010), it is an almost universal assumption in the social sciences and law that judicial independence uniformly improves judicial performance (for example, Klerman and Mahoney 2005, p. 25; Dam 2006).

[5] These facts are not only largely ignored in the literature but are often obscured. And in the following instance, they are obscured by what should be the fount of information on such facts, the Judiciary Office of the UK government: "The fundamental concept of judicial independence came into being in England and Wales in 1701 with the enactment of the Act of Settlement. This statute formally recognised the principles of security of judicial tenure by establishing that . . . [judges] . . . hold office during good behaviour. . . . Before 1701 senior judges held office at the sovereign's pleasure" (Courts and Tribunals Judiciary, Independence [https://www.judiciary.uk/about-the-judiciary/the-judiciary-the-government-and-the-constitution/jud-acc-ind/independence/]). For just one of many examples in the academic literature, see Rakove (2007, p. 1063): "The Act of Settlement of 1701 marked a true milestone in the annals of Anglo-American jurisprudence. Prior to its adoption, royal judges held their places at the pleasure of the crown."

1900). *The English Reports* are the definitive set of reports on cases heard in England's high courts before 1867.

Section 4 presents fixed-effects ordinary least squares (OLS) estimates of the effect of judicial tenure arrangements (secure tenure versus insecure tenure) on the total number of citations to a given court in a given year. Several empirical strategies aid identification of the causal effect. First, the data set is a panel, which makes possible the use of year and court fixed effects. Second, in this volatile time in English history, political conflict could have simultaneously heightened the significance of legal struggles and changed the monarch's policies on the terms of judges' appointments. Therefore, the analyses include a variable that is a proxy for such legal struggles—citations by later courts to statutes passed immediately prior to the year of court decisions. Third, the regressions use proxies for the human capital of judges, which reduces omitted-variable problems.

The core results hold in robustness tests that use variations in the citing court-year combinations used to construct the dependent variable: the results hold for citations made by the same court, made by other courts, made after 1830, and made within 20 years of the original decision. Moreover, the results are virtually identical in substance when using a tenure variable that indicates whether all judges on a court have secure tenure, which captures the presence or absence of the set of arrangements that most closely captures the intentions in the Act of Settlement.

The most plausible reason for questioning the validity of the average treatment effect (ATE) estimates from an OLS model would arise from the assumption that the monarch gave secure tenure only to judges who were less adventurous in producing new law. Section 4 presents qualitative historical evidence indicating that this assumption should be rejected. Section 5 presents instrumental variables (IV) estimates that aim to remove any bias that might arise should the troubling assumption about this monarchical appointment strategy be true. Instruments proxy the amount of conflict between king and Parliament at the time that judges were appointed. In normal times, the monarch's preferences over terms of judicial appointments would have been most important, and therefore the ATE estimates likely reflect the decisions of judges whose selection for secure tenure was relatively more influenced by the monarch. In contrast, the local average treatment effect (LATE) estimates from the IV model reflect the effect of secure tenure on the decisions of judges whose selection for secure tenure was relatively more influenced by Parliament. If there is bias in these LATE estimates, it would have the opposite sign of the most plausible bias in the ATE estimates. Given that these two sets of estimates are broadly consistent, the results from the IV estimates bolster the conclusions from the fixed-effects OLS estimates.

Section 5 also examines the sensitivity of the fixed-effects OLS estimates to potential omitted-variable bias by using the method of Oster (2019). These sensitivity tests show that it is highly unlikely that any omitted-variable bias could be large enough to vitiate the general conclusion derived from the ATE estimates of Section 4.

Concluding, Section 6 considers more speculative implications of the paper's results. It presents rough calculations of how much the change in tenure arrangements slowed the accumulation of the stock of precedents and therefore the development of English case law. The estimate is that the stock grew 20 percent less in the 18th century because of the mandated secure tenure of the Act of Settlement. These findings are not only the opposite of some hallowed wisdom on English legal history but also somewhat of a surprise because the literature on modern courts has usually concluded that stronger forms of judicial tenure are positively related to higher-quality judicial decisions (Choi, Gulati, and Posner 2010; Ash and MacLeod 2015, 2018). Therefore, Section 6 concludes with a conjecture, arguing that the early power and influence of the English legal profession could have led to the different findings for 17th- and 18th-century English judges and for those in modern times.

2. The Historical Context and Legal Framework

The ensuing empirical design is set against the background of a turbulent, but decisive, period in English history. During the Tudor years (1485–1603), Parliament had cemented itself at the center of English political life. Over an even longer period, common-law judges and lawyers had developed a sophisticated body of law and a powerful set of legal institutions. In 1603, the last Tudor died without issue, and a foreign-born monarch, James I, came to the throne. For the next 85 years, the monarchs were members of the House of Stuart, who had no appreciation of the peculiarly English position of Parliament, the importance of the common law, and the country's dominant religion. Constant political struggle ensued, which included civil wars, the beheading of a monarch, the fleeing of another, an interregnum, and the imprisonment of judges, a king, and members of Parliament.

The intensity of the political struggle abated after 1688, when Parliament reached an uneasy compromise with another foreign-born monarch, William III (Prince of Orange, who was married to the daughter of the Stuart king who had just fled). One element of this enduring compromise was the Act of Settlement of 1701, which focused on who was eligible to be monarch but, almost as an afterthought, included the clause on judicial tenure. Importantly, the clause was not a reaction to threats to judges at the time the act was passed: in 1701, all sitting judges already had appointment arrangements consistent with those stipulated by the act.

This paper examines the decisions of judges from 1600 to 1800 in the four principal high courts of justice: the King's Bench, Court of Common Pleas, Court of Exchequer, and Court of Chancery. All four were partially courts of first instance and partially appeal courts. The first three were common-law courts, and the last was an equity court. Each court specialized in a different area of law; for example, the King's Bench focused on criminal law. But the boundaries were fluid, with each court competing for the business of the others.

All four courts were under the sway of officials and judges who were steeped in the common-law tradition. Indeed, the judges in all courts were selected from a common pool of candidates, with judges often moving between courts. Each court had a chief judge and a small number of associate judges (in the terminology of the times, puisne judges).[6] The number of associate judges varied both over time within courts and across courts within time periods.

The institutional and organizational arrangements that surrounded the courts were complex and certainly not laid down in clear legislation or regulations. Even the relevant clause of the Act of Settlement was vague as to its domain of applicability. Working arrangements depended as much on informal norms as specific legislation. Perhaps for this reason, there seems to be no suitable reference that can provide a convenient summary of the institutional arrangements relevant to this paper. To fill this lacuna, Online Appendix OA compiles facts on the parts of the courts' arrangements that are useful in interpreting both the empirical design and the results that follow. The remainder of this section provides the essential points.

Before the Act of Settlement was in force, the monarch could choose to appoint judges either with secure tenure or with insecure tenure. The choice of the terms of appointments varied over time and across courts until 1714, when the judicial clause of the Act of Settlement became effective.[7] Appointments did not uniformly carry insecure tenure before 1714. Various monarchs, sometimes in reaction to political pressure, made different choices regarding the terms of judicial tenure. Their choices also varied across the four courts at a given time.

After the act was in force, the monarch had to use secure tenure for judges on the country's highest courts. In a number of ways, the act strengthened the security of tenure for judges, which is universally assumed to be a precondition of judicial independence.[8] First, appointments could no longer be insecure. Even before 1714, the conditions of secure tenure had been nearly universally respected by the monarchs if it had originally been bestowed. No doubt this was because legally removing a judge with secure tenure involved filing suit in court and hence leaving the court to decide on the merits.[9] Second, the effect of the act was not simply to transfer removal powers from monarch to Parliament. Rather, it increased veto powers over removal. The new procedures for removal of a judge

[6] The official names varied between courts. For simplicity, this paper uses the designations chief and associate only.

[7] The driving force behind the passage of the Act of Settlement was concern over who would become monarch when Queen Anne died. Thus, most clauses of the act were to become effective only when that happened, which turned out to be 1714.

[8] Before 1761, secure tenure meant holding office until retirement or death, the monarch's death, or violation of good behavior. Lifetime tenure did not arrive until 1761. All appointments expired on the death of a monarch until an act in that year.

[9] There were two cases, in 1628 and 1672, in which kings tried to remove judges appointed under secure tenure, and in both cases the pertinent judges refused to leave office because no suit had been filed in court. Both cases led to stalemate, with the judges continuing in office for a short time but unable to perform their duties in court. Both kings, having tried this stratagem once, learned their lesson and did not try it again. Instead they moved to appointing new judges at the king's pleasure.

were that there had to be an address from members of both houses requesting removal. Then the monarch had the option to remove the judge. Thus, the king and both houses of Parliament had to agree on removal. Third, the form of request by Parliament, the address, was regarded as a very solemn, high-stakes political move, unlikely to be used except in dire circumstances. In the time period covered by the data, there was no such request by Parliament that was relevant to judges. Not surprisingly then, whereas the removal of a judge from office regularly occurred in the 17th century, it was extremely rare in the 18th century. (See Online Appendix Table OA2 for further details.)

There was an important exception to the act's tenure clause: it was never interpreted as applying to the lord chancellor, the political head of the legal system and the Court of Chancery's chief judge. The lord chancellor always served at the king's pleasure.[10]

The legal system was under the sway of the legal profession, which on a number of occasions showed that it could successfully resist the demands of the monarch.[11] The monarch was not a free agent in choosing judges. They were selected from a small pool of lawyers who had risen to the top of the profession. Sitting judges and other influential members of the legal profession advised the monarch on the quality of judicial candidates. Destroying this structure would have sent the legal system crashing, and by 1600 the country was highly dependent on the functioning of that system.

Secure tenure did not mean that judges were free from acts of kingly caprice. Until the latter part of the 18th century, judges were paid from the revenues of the royal household, not the state budget. The lord chancellor, a political appointee as well as a lawyer, controlled administrative arrangements, which could have a large effect on judges' well-being.

In sum, the Act of Settlement was a legislative watershed. But it is a mistake to assume that English judges were completely dependent on the monarch before the act. And it is a mistake to assume that after the act the status of an English judge matched the modern ideal of judicial independence. Hence, the hypothesis that the Act of Settlement improved the functioning of the judiciary is just a theory. Like any theory, it should be subjected to empirical testing.

3. The Empirical Framework, the Data, and the Variables

The English Reports provides the most comprehensive record available of the decisions made by England's high courts before 1860. The empirical analysis uses a database of all citations made in the *Reports* to cases decided between 1600 and 1800 in one of England's four major courts, the King's Bench, Court of Common

[10] Note also that the act's vague wording was interpreted as not applying to courts below the country's highest level.

[11] It also could resist the demands of the head of state during the Interregnum (1649–60).

Pleas, Court of Exchequer, and Court of Chancery.[12,13] The citation database contains nearly 400,000 records, each naming the court and year of the citing case and the court and year of the cited case. Online Appendix OB provides the details of the construction of this database and the database containing the biographies of judges.

The measure of judicial performance used here is the number of citations to the reports of cases in which judges participated, a standard measure of the quality of judicial decisions in economics, law, and related fields (see Landes and Posner 1976; Landes, Lessig, and Solimine 1998; Posner 2000; Choi, Gulati, and Posner 2010; Cross and Spriggs 2010; Black and Spriggs 2013; Ash and MacLeod 2015, 2018; Fix and Fairbanks 2020).[14] Judges cite past cases to pin down the authority underlying a decision and therefore almost intrinsically point to a case with authority. If a case report is not cited, then it has not provided useful input into subsequent legal reasoning. A case report of an incoherent or bad decision is not likely to be cited for long, as it will not provide courts with compelling information to justify a decision. Good opinions are therefore cited, and bad ones are usually not cited. Thus, good opinions are an important measure of judges' quality.

Of course, the studies cited above are on the workings of the modern legal system, not on judicial behavior centuries ago. However, the core features of the modern legal system on which these studies depend is a direct product of the logic developed in English courts in much earlier times. The practice of invoking past cases was a feature of dispute adjudication in English courts as early as the 13th century and continually increased over time. Plowden and Coke, the producers of early editions of *The English Reports,* are acknowledged as providing influential examples of precedent-based logic (see Winfield 1925; Baker 2019). This approach to justifying legal decisions increased in importance monotonically thereafter, in the end solidifying in the modern notion of stare decisis.

Perhaps the most persistent criticism of the use of citations as a measure of judges' quality in the literature on modern courts is that a citation can be a negative one, critical of a previous case. In response, many scholars (for example, Cross and Spriggs 2010) point out that negative citations are quite rare and can even indicate influence—bad reasoning in past cases can be simply ignored, while

[12] This study examines time periods both before and after the Act of Settlement to produce precise estimates. Secure tenure is highly correlated with court and year. Citations are highly correlated with court and year for many reasons other than the ones in which the present study is interested. Therefore, any causal estimate of tenure effects must, at a minimum, include court and year fixed effects. In the postact period, after the fixed effects are removed from the tenure variable, there is little variation left in the key explanatory variables.

[13] A number of other courts functioned at times during these 2 centuries usually in very specific areas, such as the law of the sea or of religion. The four courts studied here were the dominant superior courts throughout 1600–1800. The highest court was Parliament, but its involvement in the legal system as a court of appeal was not routine in this period.

[14] Choi and Gulati (2007) provide an extensive list of the uses of citations as a measure of quality in the law, in the social sciences, and in the wider media. Such use has become more extensive since then. Cross and Spriggs (2010, p. 502) provide a comprehensive overview of the advantages and problems of using citations to measure the best judges and opinions, concluding that "[c]itations are the central metric for assessing the significance of opinions, at least from a legal perspective."

a negative citation is needed only against a strong past argument. The same is true of *The English Reports*. In general, it was simply not the style of the judges to explicitly mention cases with which they disagreed, instead of distinguishing which past cases applied to the issues under consideration. To investigate the prevalence of negative citations in the database derived from *The English Reports,* I examine a .1 percent sample of the 122,122 reports of cases used to derive the data for this study.[15] In the 122 sampled cases, 84 had one or more citations, with 14 of these cases containing negative citations. There were 698 citations in the sample of cases, of which only 20 citations could be interpreted as negative. Indeed, even in the 20 negative cases identified, the adjective "negative" is perhaps too strong to capture the tenor of the report.[16] On the basis of this sample, one can be confident that the deletion of negative citations would be very unlikely to have any effect on the overall set of findings of this paper.

The dependent variable in all regressions is Cite_{kt}^{m}, the number of citations to court k's decisions in year t that were made in cases decided in the court-year combinations included in the set denoted m. The term k takes on the values King's Bench, Court of Common Pleas, Court of Exchequer, or Court of Chancery; t varies from 1600 to 1800. In the most comprehensive version of Cite_{kt}^{m}, m contains all court-year combinations. Then, robustness exercises use variations in m, affording the opportunity to examine how results change when examining citations from specific courts or from specific time windows. Variations in m are introduced as they become relevant.

The focus is on the estimates of parameters β_{c} and β_{a} in the following equation:

$$\text{Cite}_{kt}^{m} = \beta_{\text{c}} C_{kt} + \beta_{\text{a}} A_{kt} + \boldsymbol{X}_{kt}\boldsymbol{\omega} + \lambda \text{Statcite}_{kt} + \phi_k + \gamma_t + \varepsilon_{kt}, \qquad (1)$$

where ϕ_k is a full set of court fixed effects; γ_t is a full set of year fixed effects; β_{c}, β_{a}, and λ are scalar parameters; $\boldsymbol{\omega}$ is a parameter vector; and ε_{kt} is a disturbance term. Table OB1 in the Online Appendix provides summary statistics.

The explanatory variables of interest are the judge-tenure measures C_{kt} and A_{kt}, the former for chief judges and the latter for associate judges; C_{kt} reflects the tenure of the chief judge in court k at time t and is a dummy variable equal to one

[15] The details of the sample and the data obtained therefrom are available from the author on request.

[16] For example, see the Chancery case *Mackenzie v. Robinson* ([1747] (3 Atkyns 559–560, 26 Eng. Rep. 1122), in which there are four positive citations to *English Report* cases and one that is ambiguously negative: "A petition was presented on behalf of a mortgagor, that the mortgagee of a naked advowson may accept of his nominee, and present him upon an avoidance. . . . Mr. *Clarke,* of counsel for the mortgagee, insisted, as there is a large arrear of interest, he ought to present . . . and cited the case of *Gardiner* versus *Griffith,* 2 *Wms.* 404 [where] the court gave no opinion, but seemed to incline that the defendant *Griffith,* the mortgagee, had a right to present. . . . *Lord Chancellor*. I am of opinion that the mortgagor ought to nominate, and it is not presumed any pecuniary advantage is made of a presentation. . . . (*Vide Amhurst* v. *Dawling,* 2 *Vern*. 401. *Attorney General* v. *Scarisbrick, ibid*. 550. *Jory* v. *Cox, Prec. Cha.* 71. *Galley* v. *Serjeant Selby,* 1 *Stra*. 403. *Com*. 343, S. C.). *Lord Chancellor* mentioned the next day, that he was not quite clear as to this point, and that he had looked into the case of *Gardiner* versus *Griffith* since yesterday . . . : he said, that was a mixed case, and that he doubted himself whether a covenant that the mortgagee should present (as was the case there) was not void."

if the appointment is with secure tenure and zero if the appointment is with insecure tenure. Because the number of associate judges varies over time and over courts and because citations are attributable only to courts, not to individual judges, the construction of A_{kt} differs slightly from that of C_{kt}. For each associate judge, a tenure dummy variable equals one if the appointment has secure tenure and zero otherwise. Then A_{kt} is the mean of the tenure dummy variables for all associate judges on court k at time t.

The panel structure admits a difference-in-differences framework. The court fixed effects (ϕ_k) absorb all time-invariant omitted variables that might induce covariation across courts between the amount of court activity and judge tenure. The year fixed effects (γ_t) prevent bias arising from time-invariant omitted variables that might induce covariation over time between courts' activity and judges' tenure.[17]

A vector of seven variables, $\boldsymbol{X}_{kt}$, proxies human capital to address the concern that the type of tenure would likely be correlated with the human capital of judges; $\boldsymbol{X}_{kt}$ includes measures capturing judges' experience and age (Choi, Gulati, and Posner 2010; Ash and MacLeod 2018).[18] As in the case of the tenure variables, the chief-judge variables in $\boldsymbol{X}_{kt}$ reflect the data for a single individual, whereas the associate-judge variables in $\boldsymbol{X}_{kt}$ reflect mean values for all associate judges on court k at time t. The variable Chief Judge Previous Service measures the number of years that the chief judge served as a judge prior to taking the current position. The variable Associate Judge Previous Service is defined analogously. The term Chief Judge Served on Same Court is a dummy variable indicating whether the chief judge previously served as an associate judge on the same court. The age variables are measured in years. The variable Chief Judge Years in Position measures the number of years that the chief judge has served in the current position at time t. The variable Associate Judges Years in Position is defined analogously.

One threat to identification of the causal effect of tenure arrangements might be that the monarch determined tenure status on the basis of expectations about the importance of litigation that was likely to come before a particular court.[19] Possibly, when that litigation was expected to raise politically controversial issues, judges in those courts would have been less likely to have appointments with secure tenure. At the same time, decisions on politically controversial cases would possibly garner more citations than other cases.

To counter this potential problem, equation (1) includes Statcite_{kt}, a term for same-court citations to statutes that measures the number of citations made after

[17] Because very few judges served under different types of tenures at different times, it is not possible to use judge fixed effects as in Ash and MacLeod (2018).

[18] Age is chosen to proxy human capital in many areas of the social science literature. However, it does not have much explanatory power for judges in modern courts (Teitelbaum 2006). The results below confirm this for 17th- and 18th-century England.

[19] Given the inclusion of fixed effects, there is no threat to identification arising from expectations about the importance of future litigation coming to courts in general.

t by court k to statutes passed in the years $t - 10$ to $t - 1$.[20] In this era, the most important political struggles were between the king and Parliament, and they often resulted in the passage of legislation that was a compromise, providing fodder for litigation in subsequent years and possibly a changed attitude to judicial appointment.[21] Therefore, it is prudent to control for expectations at time t about the likely importance of politically controversial litigation to come before court k in future years; the variable Statcite_{kt} is a proxy for such expectations. Its inclusion should diminish the potential omitted-variable problem noted just above. As it happens, the first-stage IV results presented in Section 5 show no evidence that the problem is present. Thus, the value of the inclusion of Statcite_{kt} is simply to add precision to the estimates.

4. Estimates of the Effect of Secure Tenure

This section presents fixed-effects OLS estimates of equation (1). The discussion in Section 4.1 centers on estimates of the parameters β_c and β_a, the coefficients capturing the effect of secure tenure for each type of judge. Section 4.2 presents estimates of an alternative version of equation (1), using a tenure variable that captures whether all judges on the court have secure tenure. Section 4.3 presents historical and empirical evidence arguing that the results of Sections 4.1 and 4.2 should be viewed as capturing causal estimates of the effect of secure tenure. Section 5 provides additional evidence in support of that argument, offering IV estimates that match the results of Sections 4.1 and 4.2 and using the methodology of Oster (2019) to probe the sensitivity of the OLS estimates to omitted-variable bias.

4.1. Estimating β_c and β_a

Table 1 presents fixed-effects OLS estimates of equation (1).[22] The estimation is standard; that is, it treats γ_t as a set of nuisance parameters that are solved out and not estimated. Consistency relies on a strict exogeneity assumption on the ex-

[20] Note that Statcite_{kt} varies across courts because it reflects only citations to statutes made by citing court k.

[21] Consider the Test Act of 1672, formally entitled An Act for Preventing Dangers Which May Happen from Popish Recusants (Raithby 1819, pp. 782–85). Since the king's brother was one of the recusants, the political controversy is evident. This measure inspired much important litigation, mostly confined to the King's Bench. Around this time, Charles II also began appointing judges on insecure tenure after previously always giving secure tenure. But this new policy on appointments was applied across all courts, which, after the application of fixed effects, would imply that there would be no correlation between the tenure variables and the expectation of the types of litigation coming in front of particular courts.

[22] For the sake of brevity, the results of regressions without fixed effects are not presented. As it happens, these results provide a picture that is consistent with what follows and therefore could be regarded as providing a further robustness test.

Table 1

The Effects of Chief Judges' and Associate Judges' Secure Tenure on Citations

	All Courts and Years		All Courts after 1830		All Courts within 20 Years of Decision		Same Court, All Years		Other Courts, All Years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Chief Judge Secure Tenure	66.68	39.08	14.72	2.751	40.80*	32.36+	58.66+	51.34	8.015	−12.26
	(1.50)	(.84)	(1.30)	(.23)	(2.52)	(1.84)	(1.68)	(1.48)	(.74)	(−.92)
Associate Judge Secure Tenure	−329.5***	−287.0***	−62.37***	−47.45***	−71.89***	−67.90***	−244.5***	−222.9***	−85.09***	−64.15***
	(−6.86)	(−6.04)	(−5.61)	(−3.50)	(−5.03)	(−4.53)	(−6.46)	(−6.16)	(−7.59)	(−4.80)
Chief Judge Previous Service		4.715+		.118		1.198		4.142*		.573
		(1.70)		(.16)		(1.10)		(1.99)		(.67)
Associate Judge Previous Service		14.43*		2.713		6.112*		11.16*		3.274
		(1.99)		(1.37)		(2.45)		(2.21)		(1.30)
Chief Judge Served on Same Court		−211.0***		−59.47***		−36.77*		−128.5***		−82.48***
		(−4.08)		(−4.05)		(−2.12)		(−3.36)		(−5.32)
Chief Judge Years in Position		−.334		.565		.861		−1.324		.990
		(−.15)		(.82)		(1.05)		(−.79)		(1.14)
Associate Judge Years in Position		−.629		1.773		−1.203		−3.800		3.171*
		(−.16)		(1.57)		(−.83)		(−1.34)		(2.35)
Chief Judge Age		1.543		1.058*		.215		.912		.630
		(.88)		(2.41)		(.34)		(.68)		(1.03)
Associate Judge Mean Age		1.209		−2.153*		.242		3.868+		−2.659*
		(.37)		(−2.33)		(.20)		(1.67)		(−2.34)
Statcite		.348**		.036		.101*		.301***		.0466
		(3.16)		(1.36)		(2.55)		(3.63)		(1.61)
R^2	.56	.62	.58	.63	.46	.50	.52	.59	.57	.62

Note. All regression include court and year fixed effects. The *t*-statistics are in parentheses and use standard errors clustered at the year level. $N = 804$.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

planatory variables (Wooldridge 2010, chap. 10). All estimates of standard errors are robust to heteroskedasticity and are clustered at the year level.[23]

Focus first on column 2 of Table 1, where (defining m) the citations included in the dependent variable are from all courts in all years. The key results are the estimates of the coefficients of the tenure variables β_c and β_a. For associate judges, secure tenure has a statistically significant, very large, negative effect on citations. A court that has all associate judges appointed with secure tenure will garner 287 fewer yearly citations (.89 of a standard deviation) than a court with all associate judges having insecure tenure. The regression in column 1, without any of the control variables apart from fixed effects, has an estimate of β_a that indicates a slightly stronger negative effect.

The estimated effects of secure tenure for chief judges are small, positive, and statistically insignificant in columns 1 and 2 of Table 1.[24] Moreover, the secure-tenure effect by associate judges dwarfs that of the chief judge: in the estimates in column 1, a court in which all judges have secure tenure receives 263 fewer citations than a court in which all judges could be fired at the will of the sovereign; for column 2, it is 248 fewer. These results for chief judges are substantively important because it is standard in the English history literature to assume that mandatory secure tenure had a positive effect on the legal system. In that sense, the results for chief judges and associate judges are consistent: they do not provide support for the standard assumption in the literature.

Variations in m provide robustness exercises. Columns 3–6 of Table 1 examine whether the results are sensitive to the citing time period included in m. Columns 3 and 4 provide a check on whether the substantive conclusions are affected by any change in citing practices over time. Columns 5 and 6 provide a check on whether the results are influenced by more citing years being available for earlier time periods.

The estimated β_a in the regressions with all the controls appearing in columns 4 and 6 of Table 1 are all statistically significant but indicate smaller effect sizes than in column 2 once the scale of Cite_{kt}^{m} is taken into account: the effect sizes are .55 and .62 of a standard deviation, compared with .89 for column 2. The effect

[23] This operationalizes the assumption that disturbance terms might be correlated across courts in a given year, which reflects the temporal vicissitudes of a politically volatile era. The number of year clusters (201) is sufficiently large for reliable estimates of clustered standard errors using standard, nonbootstrap procedures. Using standard criteria, the number of court clusters (four) is too small for reliable estimates of clustered standard errors using standard procedures. Section 5 presents estimates of the statistical significance of the key parameters when clustering on both court and year, using a wild bootstrap (Roodman et al. 2019).

[24] Given that the standard errors of the estimates of β_c and β_a are of similar magnitude, the differences between the statistical significances of the two estimates are not due to differential precision in estimation. A test of equality of the coefficients is rejected at the 99 percent level. To consider whether multicollinearity could be responsible for the differences in the results for the two levels of judges, the regressions in Table 1 were rerun twice with each run containing only one of the C_{kt} and A_{kt} variables. The estimates for the A_{kt} coefficients closely matched those in Table 1. The estimates for the C_{kt} coefficients were negative, economically important, and statistically significant, although still much smaller in magnitude than the corresponding coefficients for A_{kt}. This suggests some role for multicollinearity in the differences between the results for chief and associate judges. However, none of the conclusions drawn in this paper depend on those differences.

sizes for β_c are .12, .03, and .29 of a standard deviation in columns 2, 4, and 6, respectively. In all the regressions that include the controls, the estimated coefficient on β_c in column 6 is the only estimate of either β_c or β_a that is positive and significant (albeit at only the 90 percent level). In column 6, the associate-judge coefficient is twice the size of the chief-judge coefficient.

Columns 7–10 examine whether the results are sensitive to the courts used in *m*. The results in column 8—for same-court citations—provide strong support for the conclusion that secure tenure had a negative effect on judge quality. (Effect sizes are .96 and .22 of a standard deviation on β_a and β_c, respectively.) However, in the literature on modern courts, citations by courts other than the one being cited are usually regarded as the best measure of the quality of judicial decisions: within-court citations are often made routinely, whereas citations by different courts reflect the extra degree of persuasiveness needed to cross jurisdictional and subject-matter boundaries (Landes, Lessig, and Solimine 1998; Posner 2000; Choi, Gulati, and Posner 2010; Ash and MacLeod 2015, 2018). The relevant results are reported in columns 9 and 10. In the regression with all controls, the highly statistically significant estimated β_a indicates an effect size of .58 of a standard deviation, and there is a small (.1 of a standard deviation), statistically insignificant estimated negative effect for chief judges.

A summary of the results on the tenure variables would place them in the context of the existing consensus in the historical literature, that is, the theory that the Act of Settlement had a positive effect on the quality of judges' decisions. A cautious summary would be that chief-judge secure tenure has a small, statistically insignificant positive effect, and associate-judge secure tenure has a substantively large, statistically significant negative effect. There is no support for the conventional wisdom on the effect of the Act of Settlement.

The results in Table 1 for Statcite_{kt} and $\boldsymbol{X}_{kt}$ are of less direct substantive interest here. The coefficient on Statcite_{kt} is positive and statistically significant in three of the five regressions. This justifies its inclusion as a proxy for expectations of the likely importance of politically controversial litigation to come before a court. The proxies for human capital in $\boldsymbol{X}_{kt}$ have estimated effects that are most often of weak significance, statistically and substantively, perhaps because of multicollinearity. The estimated coefficients for previous years of service are sometimes statistically significant and all positive as one would expect of proxies for human capital. The variation in the estimated coefficients on the age and years-in-position variables is sufficiently large that the estimates do not offer insights. The estimated effect of the chief judge having served previously on the same court is consistently negative and statistically significant. This could be simply because of the fact that judges produce their most citable innovations when they are new to the type of cases passing through a court.

4.2. Estimating the Effects of Secure Tenure for the Whole Court

One objection to the previous framework as giving insight into the effects of the Act of Settlement is that the act dictated that all judges on specific courts

should have secure tenure, whereas the estimates of β_c and β_a capture the effects of marginal increases in secure tenure. It might be that the institution of secure tenure has a powerful effect only when all judges have secure tenure. This objection can be examined by estimating the following equation:

$$\text{Cite}_{kt}^{m} = \beta_w W_{kt} + \boldsymbol{X}_{kt}\boldsymbol{\omega} + \lambda \text{Statcite}_{kt} + \phi_k + \gamma_t + \varepsilon_{kt}, \quad (2)$$

where the only change from equation (1) is that C_{kt} and A_{kt} have been replaced by W_{kt}, a dummy variable equal to one if and only if all judges on the court have secure tenure.

The results appear in Table 2. All estimated values for β_w are statistically and substantively significant and show a larger effect than the corresponding estimates of $\beta_c + \beta_a$ derived from Table 1. This suggests that there is an extra effect due to all judges having secure tenure. However, the extra effect has the same sign as in previous estimates, which underscores previous conclusions.

4.3. Are the Estimated Effects of Secure Tenure Causal?

Acceptance of the above results as providing evidence on the causal effect of secure tenure on citations depends critically on the credibility of one assumption. This is that selection for secure tenure is conditionally independent of predictions at the time of appointment of future judges' quality as measured by later citations.

The most obvious theory questioning this assumption is that the monarch might be more likely to give secure tenure to judges who would favor the king and that those types of judges would be of a conservative disposition, less likely to produce the type of new law that would be highly cited in later years. This has surface plausibility, but it is much less than a foregone conclusion. Six reasons argue against this conclusion. First, it confounds political conservatism with legal conservatism. On a number of occasions, it was the monarch who favored positions that were legally more radical than those held by lawyers who sided with Parliament (Edie 1985). Second, judges were chosen from a narrow pool of candidates who had undergone the arduous and intellectually demanding process of rising through the legal profession. Everybody in that pool had potential, at least, to generate high-quality legal decisions.

Third, attachment to the king did not imply a lack of legal or intellectual prowess, as evidenced, for example, by the case of Francis Bacon (Coquillette 1992). Even the notorious Judge George Jeffreys, whose behavior during trials was characterized by Pound (1928, p. 11) as leaving "an abiding impression of unnecessary brutality of manner" and an "obvious and vehement partisanship," left a legacy of contributions to the law:

> In that brilliant Whig apologia, Macaulay's *History of England,* Jeffreys is portrayed as a drunken brute, quite ignorant of law, who . . . reached high judicial position as a ready instrument of the tyranny and rapacity of James II. I well remember the shock when as a first-year student of law I came upon a decision of Jeffreys as Chancellor in a collection of authorities on the law of Property. With my mind full of Macaulay's invective, it seemed

Table 2

The Effects of Whole-Court Secure Tenure on Citations

	All Courts and Years (1)	All Courts after 1830 (2)	All Courts within 20 Years of Decision (3)	Same Court, All Years (4)	Other Courts, All Years (5)
Whole Court Secure Tenure	−250.1***	−45.91***	−55.87***	−179.8***	−70.30***
	(−6.92)	(−4.46)	(−4.77)	(−6.70)	(−6.48)
Chief Judge Previous Service	4.382+	.0786	1.111	3.836*	.546
	(1.70)	(.11)	(1.09)	(2.00)	(.66)
Associate Judge Previous Service	13.54+	2.547	6.609*	10.84*	2.698
	(1.82)	(1.27)	(2.53)	(2.07)	(1.06)
Chief Judge Served on Same Court	−211.1***	−59.34***	−33.37+	−127.5**	−83.62***
	(−3.99)	(−3.92)	(−1.85)	(−3.25)	(−5.36)
Chief Judge Years in Position	−.303	.563	1.187	−1.134	.831
	(−.14)	(.85)	(1.44)	(−.70)	(.99)
Associate Judge Years in Position	.559	1.926+	.0120	−2.333	2.892*
	(.17)	(1.96)	(.01)	(−1.00)	(2.39)
Chief Judge Age	1.317	1.028*	−.108	.593	.724
	(.76)	(2.40)	(−.16)	(.45)	(1.21)
Associate Judge Mean Age	.677	−2.200**	−.512	3.053	−2.376*
	(.24)	(−2.64)	(−.52)	(1.60)	(−2.31)
Statcite	.343**	.0348	.103**	.301***	.0421
	(3.16)	(1.31)	(2.64)	(3.65)	(1.50)
R^2	.62	.63	.50	.59	.62

Note. All regression include court and year fixed effects. The *t*-statistics are in parentheses and use standard errors clustered at the year level. $N = 804$.

+ $P < .10$.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

incredible that what such a man may have decided could possibly deserve or have any authority. Later when I had to teach the law of Trusts, and hence was led to study the old equity decisions, I was astonished to find how well Jeffrey's decisions as Chancellor had maintained themselves. I was amazed to find how much more he counted in the reports which have made our law than Somers, who is next to William III, the hero of Macaulay's history. . . . Whatever else he may have been, the law books show clearly enough that Jeffreys was a lawyer—indeed was no ordinary lawyer. (Pound 1928, pp. 7–8)

Pound's amazement surely reflected his surprise at realizing that he had to reject that enticing theory that judges who favored the king were the ones who were less likely to produce innovative law.

Fourth, the results in column 4 of Table 1 and column 2 of Table 2 reflect citations made after 1830, long after the passage of time had made settled law of the issues that animated 17th-century monarchs and Parliament. Fifth, given the lackluster performance in forecasting judicial decisions by modern experts with a wealth of information (Ruger et al. 2004),[25] one might be attributing a little too much perspicacity to the monarchs in assuming that they could look into the future and make a prediction of a judge's performance in matters of law. Many of the monarchs were educated in environments that placed no emphasis on the workings of the common law, which would have diminished their forecasting capacity.[26] Certainly, James I interacted closely with the greatest lawyer of them all, Edward Coke, before appointing him chief justice of common pleas, but later James was more than a little surprised by Coke's performance (Holdsworth 1935).

Sixth, some quantitative information can be brought to bear on implications of the theory that judges favorable to the monarch would be simultaneously more likely to gain secure tenure and less likely to be of high quality. If the monarch were using predictions of the future performance of judges when selecting for secure tenure, then such selection would have been most important when the courts were likely to face particularly controversial issues. But results reported in Online Appendix OC show that there is no significant difference between the mean values of predictors of the importance of impending litigation when new judges were appointed with secure tenure and the analogous mean values for ap-

[25] In the 2002 Supreme Court term, a group of experts in the pertinent areas of the law forecast 59 percent of binary decisions correctly (Ruger et al. 2004). The hypothesis that the forecasts were no better than random was not rejected at the 1 percent significance level.

[26] James I was the monarch who was most knowledgeable about law, but to the frustration of his critics, he viewed the law through a European prism. Charles I's education was most influenced by his father and a Scottish tutor. Charles II spent the years when he might have learned about the law under the influence of his mother, a member of the French royal family. Both demand and supply assured that the education of James II was limited. William III was 39 years old before he took the English throne. His wife, Mary, and her sister, Queen Anne, received the standard amount of political and legal education for women at that time, virtually nothing. George I came to England when he was 54 and had only a smattering of English. There is no evidence that George II or George III possessed either the intellect or the preferences necessary to gain a deep understanding of British legal and constitutional matters.

pointments with insecure tenure.[27] The same result is shown for firings and nonrenewals of judges, which occurred more often when the predicted importance of impending litigation was less. Thus, in examining the circumstances surrounding monarchs' hiring and firing decisions, Online Appendix OC provides no evidence for direct implications of the theory that monarchs gave secure tenure to trusted judges who, in turn, would not produce citable decisions.[28] This evidence can be viewed as a substitute for the standard pretreatment trends analysis that is not feasible given the structure of the current data set.

All six of these arguments serve to bolster the credibility of the crucial assumption that selection for secure tenure is conditionally independent of predictions at the time of appointment of future judges' quality and therefore that the unobservable characteristics omitted from equation (1) are not likely to be important in explaining selection for treatment. Therefore, the results presented in Tables 1 and 2 arguably reflect the causal effect of secure tenure on citations.

5. Robustness Exercises

This section presents further evidence bolstering the conclusion of Section 4.3. The evidence is provided in a series of robustness exercises. None of the many details of the exercises alter in any important way the interpretation of the findings. Therefore, many of those details, including the tables of results, are relegated to the Online Appendix. Instead, the focus here is on providing an overview of procedures and results. Section 5.1 describes IV estimates. Section 5.2 summarizes an implementation of the Oster (2019) procedure that examines the sensitivity of OLS estimates to omitted-variable bias. Section 5.3 reports on bootstrap estimates of standard errors for the OLS and IV estimates, particularly by implementing clustering on courts. Section 5.4 examines sensitivity to the presence of outliers by modifying the measurement of the dependent variable.

5.1. Instrumental Variables Estimates of the Effect of Secure Tenure

As already noted, the most obvious theory predicting a bias in OLS estimates is that there is a link between the types of judges selected by the monarch and the subsequent citability of those judges' decisions. One way to diminish, or even reverse, this bias is to obtain estimates of β_c, β_a, and β_w that more closely reflect the effect of secure tenure for judges appointed at times when the monarch was relatively constrained in choosing the terms of judicial appointments. Implementation of this approach uses an instrument reflecting the tightness of those constraints on the monarch to produce LATE estimates.

[27] The mean values are calculated using residuals from regressions on court and year fixed effects, since there would be no omitted-variable bias arising from monarchs' preferences that did not vary across courts in particular years or across years in particular courts.

[28] Given the available data, it is not possible to directly test the hypothesis that monarchs were able to pick for secure tenure the types of judges who would be reliable and therefore not be highly cited. Instead, it is necessary to examine implications one step removed from the hypothesis. Details are provided in Online Appendix OC.

The defining struggle in the era surrounding the Act of Settlement was between monarch and Parliament. One element of this struggle was Parliament's preference to make the judges more independent of the monarch.[29] The supporters of Parliament would have had relatively little influence on judges' appointments when the monarch was either strong enough or determined enough not to call Parliament. In contrast, when Parliament was meeting regularly, either the differences in opinion between the monarch and Parliament were muted or the monarch needed Parliament's cooperation because the government was in need of parliamentary approval of increased financing. The required instrument captures parliamentary activity: the number of days that Parliament was in session over the 10 years previous to t. But given the inclusion of the fixed effects, an instrument must vary across k for given t. Therefore, the number of days Parliament is in session is interacted with each of the court dummy variables. This reflects the fact that king and Parliament would have been more focused on some courts than others. The two courts whose litigation was mostly likely to reflect political conflict or cooperation were the King's Bench and the Common Pleas. The excludability condition for instrument validity is satisfied because the inclusion of Statcite_{kt} in the second-stage regression captures any direct link from political conditions to citation rates; Statcite_{kt} is formulated exactly to operationalize this link.

Even though the interaction term gives four variables, only three are independent, and they reflect only two pieces of information, political conflict and relative preference over courts. Given two endogenous variables, another instrument would be useful. The process of the selection of judges in 17th- and 18th-century England was shrouded in mystery, and there is little in the historical record to explain which judges received secure tenure and which did not. It is necessary to conjecture. Perhaps the more years the chief judge was expected to serve, the less likely that new judges on the court would be given secure tenure because that would have locked in the makeup of the court for many years, an outcome that would have reduced future options for monarch and profession. The additional instrument is the expected life expectancy of the chief judge of the court, proxied by the number of years that the judge lived.[30] The excludability condition for instrument validity is satisfied because this is purely a measure of the expected longevity of a chief judge's position on the court rather than his qualities as a judge. To the extent that there are concerns that life expectancy might proxy human capital, these would be blunted by the inclusion of the seven variables in $\boldsymbol{X}_{kt}$.

It is worthwhile emphasizing what type of selection for treatment the IV estimates reflect. For the interaction term of days Parliament is in session and court dummies, the LATE estimates reflect the effect of secure tenure on the decisions of judges who have two characteristics: they would receive secure tenure when Parliament had relatively more influence over judge-appointment decisions, and

[29] Parliament let its preferences be known formally in 1640, 1645, 1674, 1680, 1689, 1691, and 1696. See Online Appendix Table OA1 for details.

[30] Of course the death date is unknown at the time of appointment, but it is the best available proxy for contemporaneous expectations concerning the future health of a potential appointee.

they would not receive secure tenure when the monarch could safely ignore Parliament. Their secure tenure reflects Parliament's preferences and not the monarch's. Therefore the IV estimates are not subject to the type of bias that would arise from a theorized link between the types of judges selected for secure tenure by the monarch and the subsequent (non)citability of those judges' decisions. Those same concerns would not arise for judges favored by Parliament.

Table OD1 in the Online Appendix presents the first-stage regressions, together with standard diagnostics, which suggest strong instruments. The estimated coefficients of the political conflict variables in the regressions indicate that in periods of greater conflict, there would be fewer appointments with secure tenure in the King's Bench and the Common Pleas. Chief judges' life expectancy has a negative effect.

Table OD2 presents the IV estimates of equation (1). Crudely summarized, the β_a values are an average of 10 percent smaller in absolute magnitude than the corresponding coefficients from the OLS model; the estimated sizes of the β_c values, on average, differ little from those of the OLS model. Naturally, the standard errors are considerably higher in IV estimation than in the OLS model, but even so a majority of the estimated β_a values are significant at least at the 5 percent level. The β_c values are all nonsignificant. Tables OD3 and OD4 provide the results for a similar exercise for whole-court secure tenure. The β_w values from the IV model are an average of 17 percent smaller in absolute magnitude than the corresponding coefficients from the OLS model. Three of the five coefficients are statistically significant at least at the 5 percent level.

The results of Sections 4 and 5 complement each other when drawing overall conclusions. The ATE estimates in Section 4 reflect the decisions of judges whose terms of appointment were relatively more influenced by the monarch. The LATE estimates in this section reflect the effect of secure tenure on the decisions of judges whose terms of appointment were relatively more influenced by Parliament. If there were bias resulting from the characteristics of the types of judges that were selected for secure tenure, then one would expect to see stark differences between these two sets of estimates. In fact, they give broadly consistent results, thereby bolstering each other in providing evidence for this paper's overall conclusions.

5.2. The Sensitivity of the Ordinary Least Squares Estimates to Omitted-Variable Bias

An alternative approach to examining the validity of the evidence from the OLS model is to probe the sensitivity of those estimates to possible omitted-variable bias. One immediate perspective can be gained from a comparison of the results within the five pairs of columns in Table 1.[31] Addition of the eight controls reduces the absolute size of the estimated β_a by only an average of 15 percent over the

[31] In the following discussion, the fixed effects are always included in the pertinent regressions and thus are not included in the set of variables referred to as controls.

five pairs, which slightly weakens the negative effect of secure tenure for associate judges. In reducing the estimated β_c by an average of 40 percent, the addition of the controls weakens the already weak positive effect of secure tenure for chief judges. The insight of Altonji, Elder, and Taber (2005) is that the size of these changes provides information on the likely effect of unobservables. Oster (2019) builds on this insight and develops a method of estimating bounds on the causal value of the coefficient of interest.

The details of the application of the Oster (2019) method are relegated to Online Appendix OE, as are the results from that application. The briefest summary suffices here. The procedure using the Altonji, Elder, and Taber (2005) and Oster (2019) methods is generally viewed as providing estimates of bounds on the likely true coefficient and produces an estimated interval in which an estimated coefficient would almost certainly lie if all omitted-variable problems were solved. Thus, for example, while the estimate of β_a reported in column 2 of Table 1 is -287, the estimated interval for β_a is $[-287, -169]$. Using three alternative assumptions to implement the Oster (2019) procedure, 15 such intervals are presented in Table OE1, three for each of the five relevant β_a estimates of Table 1 (those in even-numbered columns). Only three of the 15 intervals contain 0, and the major part of each of the three intervals lies in negative territory. It is unnecessary to carry out the same exercise for the estimates of β_c in Table 1 because the estimates of β_c with controls are all smaller than the estimates without controls. Application of the Oster (2019) procedure to these β_c would therefore lead to an interval whose maximum is the estimated β_c and whose minimum is closer to 0 or, more likely, negative, thereby providing evidence consistent with previous conclusions.

When applying the Oster (2019) procedure to the β_w estimates reported in Table 2, the overall conclusion is nearly identical to that reached for the analogous exercise on β_a (see Table OE2). The almost universal hypothesis in the historical literature is that β_c, β_a, and β_w should all be positive, and the application of this procedure provides further evidence that this hypothesis should be rejected.

5.3. Standard Errors When Clustering on Courts and Years

The results reported so far use estimates of standard errors assuming clustering on years. The number of year clusters (201) is sufficient for large-sample theory to provide reasonable criteria for interpretation of the estimates, which facilitates the use of standard techniques. However, the number of court clusters (four) is sufficiently small as to necessitate the use of alternative techniques. The wild bootstrap has been found to perform well in such circumstances (Roodman et al. 2019). Online Appendix OF presents *p*-values and confidence sets for all estimates of β_c, β_a, and β_w reported above using a wild bootstrap with clustering on both court and year.

The standard errors of coefficients reported in Online Appendix OF are larger than the corresponding estimates in Tables 1 and 2, as is usual. However, the pat-

tern of previous results on statistical significance is preserved. All of the β_a values reported in Table OF1 corresponding to those in Table 1 are statistically significant at the 10 percent level when there is clustering on both court and year. None of the β_c values in Table OF1 are statistically significant at the 10 percent level. Four of the five β_w values reported in Table OF2 are statistically significant at the 1 percent level, while the other is significant at the 5 percent level. There is very little change in the significance levels of the estimated β_c and β_a values produced by the IV method. (Compare Tables OD2 and OD4 to Tables OF3 and OF4.)

5.4. Using the Logarithm of Citations as the Dependent Variable

Online Appendix OG provides a further set of robustness exercises using $\ln(1 + \text{Cite}_{kt}^{m})$ instead of Cite_{kt}^{m} as the dependent variable. This transformation provides a check on whether estimates are unduly affected by large positive outliers in the dependent variable. As it happens, the results provide even stronger support for the main conclusions of this paper.[32]

6. Conclusions and Conjectures

This paper has examined the effect of judicial tenure arrangements, and implicitly the Act of Settlement, on the quality of judicial decision-making during England's formative years of modern development. Conferring secure tenure resulted in a large decline in the quality of the decisions of associate judges, while perhaps improving the quality of the decisions of chief judges, but the net effect for the court as a whole was negative. In a common-law system, judges' decisions establish precedents, the reasons for applying legal rules that are implicitly used in decisions.[33] Then the body of precedents forms a legal capital stock that later supplies information for judges or private agents. One measure of the amount of legal development is the growth of this precedential capital stock, with the importance of the information flows from specific precedents measured by the number of citations to them. From this perspective, the judicial clause of the Act of Settlement delayed the development of English law.

By how much was development slowed? Using the estimates in column 2 of Table 1 and assuming that tenuring rates from 1715 to 1800 were at the mean rate for the period 1600–1714, the counterfactual prediction is that decisions made from 1715 to 1800 were cited at a rate 19.5 percent lower than they would have been had the Act of Settlement not been in force. This is obviously a very crude measure of the amount of precedential reasoning lost, but it is one indication of the substantive significance of the estimates reported above.

The overall conclusion of this paper is at odds with the conventional wisdom

[32] In a further robustness test addressing the issue of the effect of outliers, all observations from the data set that contained more than 50 citations for the dependent variable were deleted. Then all the regressions in the paper were rerun. When comparing the results for the restricted data set and those in Tables 1, 2, OD2, and OD4, one can conclude that none of the paper's core conclusions would be changed by relying on the restricted data set. These results are available from the author on request.

[33] This formulation follows closely that of Landes and Posner (1976).

that has held sway for more than 2 centuries. To quote Justice John Marshall Harlan from 1891, expressing a sentiment that reaches back to Hamilton and forward to today,[34] "No one, in my judgment, under our system of law, can be appointed a judge . . . to hold the office at the pleasure and will of another. No such doctrine has been maintained in England since the [Act of Settlement,] . . . one of the great acts which followed the revolution of 1688. Previously to that period most of the judges of the higher courts held their offices during the pleasure of the crown. . . . This power exerted a most baleful influence" (*McAllister v. United States*, 141 U.S. 174, 187 [1891]).

Are this paper's results externally generalizable beyond this historical episode? Certainly, there are several studies on modern courts that reach the opposite conclusion on the effect of different forms of tenure on the quality of judges' decisions (Choi, Gulati, and Posner 2010; Ash and MacLeod 2015, 2018). Perhaps this paper's most important lessons are confined to a particular period of institutional and political development. However, this period saw institutional changes of profound importance, which Hamilton, Harlan, and many of today's scholars regard as providing paradigmatic lessons about institutions, growth, and development (for example, North and Weingast 1989; Olson 1993; Acemoglu and Robinson 2012; Besley and Ghatak 2010). It is in this sense that the lessons of this paper apply to a much broader canvas than that of English history. Not least among these lessons is that the effect of institutional changes can be highly context contingent. And given that 17th-century England was a developing economy, perhaps the present paper's results are germane when deliberating on today's developing countries.

What was the context of 17th- and 18th-century England that could have led to the negative effect where a positive one has been the universal prior belief? This question is beyond the scope of this paper's empirical analysis. But one conjecture reflects the sentiment in the quote from Justice Sandra Day O'Connor in one of this paper's epigraphs: incentives. The legal profession was very powerful.[35] Untenured judges had two principals, officially the monarch and unofficially their profession. The monarch would have been reviewing for political reliability, but the profession was reviewing for quality. The admiration of the legal profession was the judge's best protection when facing an aggrieved monarch in a very dangerous political environment. The protection would have been greater for judges perceived by their peers as producing higher-quality decisions.

These speculations gain force when comparing the results for chief judges and

[34] Studies moving forward through time illustrate this. McIlwain (1913, p. 217) states that "[i]n the history of the tenure of English judges the . . . Act of Settlement, is the greatest landmark"; Tarkow (1943) emphasizes the tenure clause while comparing the Act of Settlement with the Magna Carta; Shetreet and Turenne (2013) view the Act of Settlement as fundamental in securing judicial independence, which they regard as essential to the rule of law; and Baker (2019, pp. 178–79) comments that contemporaries believed that the best solution to the problem of Crown intervention into the courts was the awarding of life tenure for judges.

[35] On the rise and the strength of the English legal professions, see, for example, Baker (1986), Cromartie (2006), Plucknett (1983), and Halliday and Karpik (1997). One measure of that strength is that in the first half of the 17th century over one-half of the members of Parliament had been educated in the Inns of Court (Cromartie 2006, pp. 180–81).

associate judges, the effect of secure tenure being clearly negative for the latter and positive, but much closer to 0, for the former. Such was the prestige of the law in England that the chief judges were politically powerful simply by dint of their positions. They had already risen to such a high level of prestige that the protection of the profession probably did not mean much for them: they were the ones who provided the political protection for lesser members of the profession. This could explain why the effect of secure tenure was insignificant for chief judges.

The situation was certainly different for associate judges, who might be especially vulnerable. Under secure tenure, associate judges could treat their position as a sinecure if they did not have either intrinsic motivation or ambition to rise in the legal hierarchy. Arguably, they worked less hard or their decisions became tainted by the many opportunities available at that time to make money out of a judicial position. There were very lax rules on what constituted corruption and what was a conflict of interest, which made outside sources of income extremely lucrative for the judges. Indeed, in the time period covered by the data, the amount of money to be earned from judgeships and practice of the law in general was as high, relatively speaking, as at any time in English history (Rubinstein 1983; Wasson 1998). Notably, when the chief judges advised the king in 1691 on a bill that contained features that were similar to those that were eventually embodied in the judicial clause of the Act of Settlement, they encouraged him not to sign because they thought it was not appropriate for a judge to be independent of the Crown (Burnet 1734, p. 86).[36] A chief judge would find it much harder to control an independent associate judge.

This reasoning helps to explain the differences between this paper's conclusions and those in the studies on modern US courts that focus on the effects of different forms of tenure on the quality of judicial decisions. Judges are the agents. The principals differ under alternative institutional arrangements. In the analyses of modern courts, judges who face competitive elections have insecure tenure, with the public and politicians becoming important principals. The absence of such elections gives securer tenure, which strengthens the relative importance of the incentives arising from intrinsic motivation and reputation within the broader legal profession (Posner 1993). This incentivizes higher-quality decisions.[37] In the modern era then, secure tenure increases the relative importance of the incentives to produce decisions that are recognized as having high quality.

In 17th- and 18th-century England, judges with insecure tenure implicitly

[36] The bill to which Burnet was referring was entitled An Act for Ascertaining the Commissions and Salaries of Judges. Burnet (1734, p. 86) wrote, "Among the Bills that were offered to the King, at the end of the Session, one was to secure the Judges Salaries; and to put it out of the King's power to stop them. The Judges had their Commission, during their good behaviour; Yet their Salaries were not so secured to them, but that these were at the King's pleasure. But the King put a stop to this, and refused to pass the Bill: for it was represented to him, by some of the Judges themselves, that it was not fit they should be out of all dependence on the Court." I have interpreted "some of the judges" to mean the chief judges, since they would certainly have been those consulted by the king.

[37] Lim (2013) shows that judges receive large nonpecuniary benefits from their employment and that following the preferences of their principals reduces these benefits significantly.

faced a continuous review of performance and would, just like modern judges facing election challenges, be aware of their principals' preferences (Shepherd 2009). However, in contrast to modern times, a strengthening of tenure arrangements in 17th- and 18th-century England would have decreased the relative importance of the incentives induced by the legal profession, the principal that would have focused on the quality of decisions. In early England then, only intrinsic motivation would have increased in relative importance on the awarding of secure tenure, which would give broader scope for a Judge Roy Bean effect. Thus, in historical England, the incentives provided by the legal profession became less important after the awarding of secure tenure, in contrast to the effects of secure tenure in modern courts. There is no fundamental inconsistency in the contrasting empirical results from the modern era and those of this paper, once one considers the differing institutional contexts.[38]

References

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113:151–84.

Ash, Elliott, and W. Bentley MacLeod. 2015. Intrinsic Motivation in Public Service: Theory and Evidence from State Supreme Courts. *Journal of Law and Economics* 58:863–913.

———. 2018. Elections as Incentive and Selection Device: The Case of State Supreme Courts. Working paper. ETH Zurich, Department of Humanities, Social and Political Sciences, Zurich.

Baker, John H. 1986. *The Legal Profession and the Common Law: Historical Essays*. London: Hambledon Press.

———. 2019. *An Introduction to English Legal History*. 5th ed. Oxford: Oxford University Press.

Berger, Raoul. 1970. Impeachment of Judges and "Good Behavior" Tenure. *Yale Law Journal* 79:1475–1531.

Besley, Timothy, and Maitreesh Ghatak. 2010. Improvement and Extension of Property Rights. Pp. 5:4525–95 in *Handbook of Development Economics,* edited by Dani Rodrik and Mark Rosenzweig. Amsterdam: North-Holland.

Black, Ryan C., and James F. Spriggs II. 2013. The Citation and Depreciation of U.S. Supreme Court Precedent. *Journal of Empirical Legal Studies* 10:325–58.

Burnet, Gilbert. 1734. *Bishop Burnet's History of His Own Time: Volume II from the Restoration of King Charles II, to the Conclusion of the Treaty of Peace at Utrecht, in the Reign of Queen Anne*. London: Joseph Downing & Henry Woodfall.

Choi, Stephen J., and Mitu Gulati. 2007. Ranking Judges according to Citation Bias (as a Means to Reduce Bias). *Notre Dame Law Review* 82:1279–1309.

Choi, Stephen J., G. Mitu Gulati, and Eric A. Posner. 2010. Professionals or Politicians:

[38] One example from modern times that is consistent with the results from English history is the alignment between the objectives of the legal profession and the intrinsic objectives of judges in US federal bankruptcy courts, which McKenzie (2010) argues is the reason why secure tenure is not an issue for the currently untenured judges.

The Uncertain Empirical Case for an Elected Rather than Appointed Judiciary. *Journal of Law, Economics, and Organization* 26:290–336.
Coquillette, Daniel R. 1992. *Francis Bacon*. Stanford, CA: Stanford University Press.
Cromartie, Alan. 2006. *The Constitutionalist Revolution: An Essay on the History of England, 1450–1642*. Cambridge: Cambridge University Press.
Cross, Frank B., and James F. Spriggs II. 2010. The Most Important (and Best) Supreme Court Opinions and Justices. *Emory Law Journal* 60:407–502.
Dam, Kenneth W. 2006. *The Law-Growth Nexus: The Rule of Law and Economic Development*. Washington, DC: Brookings Institution.
Edie, Carolyn A. 1985. Tactics and Strategies: Parliament's Attack upon the Royal Dispensing Power, 1597–1689. *American Journal of Legal History* 29:197–234.
Feld, Lars P., and Stefan Voigt. 2003. Economic Growth and Judicial Independence: Cross-Country Evidence Using a New Set of Indicators. *European Journal of Political Economy* 19:497–527.
Fix, Michael P., and Bailey R. Fairbanks. 2020. The Effect of Opinion Readability on the Impact of U.S. Supreme Court Precedents in State High Courts. *Social Science Quarterly* 101:811–24.
Halliday, Terence C., and Lucien Karpik, eds. 1997. *Lawyers and the Rise of Western Political Liberalism: Europe and North America from the Eighteenth to Twentieth Centuries*. Oxford: Clarendon Press.
Hamilton, Alexander. (1788) 1981. The Federalist No. 78. Pp. 226–33 in *The Federalist Papers: A Collection of Essays Written in Support of the Constitution of the United States: From the Original Text of Alexander Hamilton, James Madison, John Jay,* edited by Roy P. Fairfield. Baltimore: Johns Hopkins University Press.
Hanssen, Andrew. 2004. Is There a Politically Optimal Level of Judicial Independence? *American Economic Review* 94:712–29.
Holdsworth, William S. 1935. Sir Edward Coke. *Cambridge Law Journal* 5:332–46.
Klerman, Daniel M., and Paul G. Mahoney. 2005. The Value of Judicial Independence: Evidence from Eighteenth Century England. *American Law and Economics Review* 7:1–27.
Landes, William M., Lawrence Lessig, and Michael E. Solimine. 1998. Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges. *Journal of Legal Studies* 27:271–332.
Landes, William M., and Richard A. Posner. 1976. Legal Precedent: A Theoretical and Empirical Analysis. *Journal of Law and Economics* 19:249–307.
Lim, Claire S. H. 2013. Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges. *American Economic Review* 103:1360–97.
McIlwain, C. H. 1913. The Tenure of English Judges. *American Political Science Review* 7:217–29.
McKenzie, Troy A. 2010. Judicial Independence, Autonomy, and the Bankruptcy Courts. *Stanford Law Review* 62:747–807.
Murrell, Peter. 2017. Judicial Tenure and the Slowing of Legal Development in England, 1600–1800. Working paper. University of Maryland, Department of Economics, College Park.
North, Douglass C., and Barry R. Weingast. 1989. Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England. *Journal of Economic History* 49:803–32.
O'Connor, Sandra Day. 2009. Judicial Independence. *Bulletin of the American Academy of Arts and Sciences* 62:46–48.
Olson, Mancur. 1993. Dictatorship, Democracy, and Development. *American Political Sci-*

ence Review 87:567–76.
Oster, Emily. 2019. Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business and Economic Statistics* 37:187–204.
Plucknett, Theodore F. T. 1983. *Studies in English Legal History*. London: Hambledon Press.
Posner, Richard A. 1993. What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does). *Supreme Court Economic Review* 3:1–41.
———. 2000. An Economic Analysis of the Use of Citations in the Law. *American Law and Economics Review* 2:381–406.
Pound, Roscoe. 1928. The American Attitude toward the Trial Judge. *Dakota Law Review* 2:5–16.
Raithby, John, ed. 1819. *Statutes of the Realm: Volume 5, 1628–80.* London: Great Britain Record Commission.
———. 1820. *Statutes of the Realm: Volume 7, 1695–1701*. London: Great Britain Record Commission.
Rakove, Jack N. 2007. The Original Justifications for Judicial Independence. *Georgetown Law Journal* 95:1061–76.
Renton, Alexander Wood, ed. 1900. *The English Reports*. Edinburgh: W. Green & Sons.
Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. Fast and Wild: Bootstrap Inference in Stata Using boottest. *Stata Journal* 19:4–60.
Rubinstein, W. D. 1983. The End of "Old Corruption" in Britain, 1780–1860. *Past and Present* 101:55–86.
Ruger, Theodore W., Pauline T. Kim, Andrew D. Martin, and Kevin M. Quinn. 2004. The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking. *Columbia Law Review* 104:1150–1209.
Shepherd, Joanna M. 2009. The Influence of Retention Politics on Judges' Voting. *Journal of Legal Studies* 38:169–206.
Shetreet, Shimon, and Sophie Turenne. 2013. *Judges on Trial: The Independence and Accountability of the English Judiciary*. 2d ed. Cambridge: Cambridge University Press.
Tarkow, I. Naamani. 1943. The Significance of the Act of Settlement in the Evolution of English Democracy. *Political Science Quarterly* 58:537–61.
Teitelbaum, Joshua C. 2006. Age and Tenure of the Justices and Productivity of the U.S. Supreme Court: Are Term Limits Necessary? *Florida State University Law Review* 34:161–81.
Wasson, E. A. 1998. The Penetration of New Wealth into the English Governing Class from the Middle Ages to the First World War. *Economic History Review* 51:25–48.
Winfield, Percy H. 1925. *The Chief Sources of English Legal History.* Cambridge, MA: Harvard University Press.
Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* 2d ed. Cambridge, MA: MIT Press.