

**Judging Political Promotion of Judges:
Survival Analysis, Split Population Model and Matching Method**

Kentaro Fukumoto^{*} and Mikitaka Masuyama^{**}

Prepared for delivery at the 2006 Annual Meeting of the American Political Science Association,
Philadelphia, August 31–September 3, 2006.

ABSTRACT

Ramseyer and Rasmusen (2003) argue that a judge who once belonged to a leftist group takes longer to reach a moderately prestigious administrative status, and that Japanese judges are conservatively biased. Their findings are, however, dependent on the assumption that they observe or completely estimate the timing of promotion for all judges, which is not correct. In an attempt to deal with the issues of censoring and time dependence, we apply survival analysis to the data set of judicial careers, which provides us with a methodological challenge to solve simultaneously the problems of left truncation and split population. Moreover, we propose a way to estimate an average treatment effect on duration time and to perform matching to reduce model dependence of our analysis. A most important finding is that leftist judges are less likely to be on the track to a prestigious post, but once they are on the track, they are not discriminated in terms of the timing of promotion.

^{*} Associate Professor, Department of Political Science, Gakushuin University, Tokyo.

E-mail: First Name dot Last Name at gakushuin dot ac dot jp.

URL: <http://www-cc.gakushuin.ac.jp/~e982440/index_e.htm>.

^{**} Professor, Department of Political Science, Keio University, Tokyo.

E-mail: Last Name at law dot keio dot ac dot jp.

URL: <<http://www.law.keio.ac.jp/~masuyama/>>.

INTRODUCTION

The Japanese Constitution stipulates that all judges shall be independent in the exercise of their conscience and shall be bound only by the Constitution and the laws. At the same time, it provides the Cabinet with the power to designate a Chief Supreme Court Judge (appointed by the Emperor) and to appoint other Supreme Court Judges. The judges of the inferior courts shall be nominated by the Supreme Court and appointed by the Cabinet, while no executive organ is empowered to administer disciplinary action against judges. Ramseyer and Rasmusen (2003) argue that, despite the Constitution, Japanese judges routinely validate what the government has done, and question the conventional view in which the Japanese judicial system is independent of political control. In order to substantiate their view, they have shown in a series of statistical examinations that judges with leftist preferences do worse in their careers. While agreeing on the importance of analyzing judicial independence, we suggest that their methods are problematic.

To be concrete, Ramseyer and Rasmusen examine whether or not a judge who once belonged to a leftist group took longer to reach the moderately prestigious administrative status known as *sokatsu*, hereafter we call a “division chief judge (DCJ)”.¹ In order to regress the time to DCJ in an OLS model, they drop some judges from the career data set and substitute arbitrary values for those who even did not reach DCJ. The problem with this data manipulation is not only to cause bias in the estimates, but also to ignore the time-dependence of an event such as a judge obtaining a DCJ appointment.

Survival analysis provides an efficient solution to simultaneously deal with the issues of censoring and time-dependence. In particular, we consider the possibility that some judges may have a chance to become DCJ, while others may have no chance at all. To account for the problem of incorrectly assuming all observations will eventually make a transition, we apply a split population model to the data set of judicial career. Also, we attempt to deal with the issue of what is called “left truncation,” an issue rarely addressed in the context of split population modeling. At the same time, we take causal inference seriously and match every observation with its similar one so that our survival analysis is less dependent on the particular models we specify.

This paper proceeds as follows. We begin with a brief review of the argument made by Ramseyer and Rasmusen (2003) and a replication of the statistical analysis reported in their book. The second section reconsiders their analysis. First, we correct their career data and report the OLS results using the corrected data set. Second, we turn to an explanation of

¹ A judge with this status is the head of the division of a court and eligible to preside court for the cases requiring consultation with associate judges.

survival analysis, left truncation and split population model, and a description of our data set, which treats the time needed for a judge to become DCJ as a duration time in a survival analysis setting. Third, we extend our model to other judicial postings that Ramseyer and Rasmusen did not consider. Fourth, we discuss the issue of causal inference and introduce a matching method. In the subsequent section, we conduct matching and survival analysis of judicial promotion. Our findings claim that leftist judges are less likely to be on the track to DCJ, although once they are on the track, their timing to DCJ is not discernible from that of other judges. We conclude with a summary of the findings and a brief discussion of their implications for a better understanding of judicial independence in Japan and elsewhere.

REPLICATION

Ramseyer and Rasmusen argue that judges with leftist preferences and/or enjoining the conservative government do worse in their careers. In their view, the Japanese courts use job postings as incentives and judges uphold the conservative positions of the longtime incumbent Liberal Democratic Party. Among many statistical analyses, Ramseyer and Rasmusen use the career data of roughly 500 judges and show that judges who joined a leftist group were promoted more slowly than their peers. In this paper, we begin by replicating the statistical analysis of judicial career reported in Ramseyer and Rasmusen (2003, p.41).

Using a list that details all posts held by judges (Nihon Minshu Horitsuka Kyokai Shiho Seido Iinkai 1998), Ramseyer and Rasmusen (2003, p.39) compiled the judicial career data for those hired between 1959 and 1968. Although it is ideal to have the pay data on individual judges, there is no such data. Instead, they focus on the time it takes for a judge to reach the status of DCJ, as a proxy for the rate at which a judge climbs the pay scale, since generally judges do not receive a DCJ posting until they reach a specific rank in the pay scale. A variable they call *Time2Sok* is defined as the year a judge first received a DCJ appointment, less the year he or she graduated from the Legal Research and Training Institute.

Ramseyer and Rasmusen (2003, pp.177-178) make two main adjustments. First, they drop those judges who held non-judicial postings (generally regarded as prestigious) in the two years before their first DCJ posting (hereafter we call these judges “stars”). In the rare cases where a judge served as chief judge before serving as DCJ, they keep such cases in the data set and treat the appointment as chief judge as a DCJ posting, since such an appointment is unambiguously higher than DCJ. Second, they drop the group of 164 judges who (1) never obtained a DCJ appointment and (2) quit or died before the mean *Time2Sok* for the rest of the group, 20.41 years (whom we call “dropouts”). However, they kept 83 judges who quit or died after 20.41 years without a DCJ appointment and treated their death or resignation as their time

of first DCJ appointment (whom we call “losers”). Also, they include the judges who did not have a DCJ appointment as of 1997, the last year the list of judges reports career data. In addition to these adjustments, they drop judges who did not join the court within a year of graduating from the Legal Research and Training Institute. Of 797 judges listed for the period between 1959 and 1968, 501 remain in the career data set.

To test for political discrimination, Ramseyer and Rasmusen introduce a variable named *YJL*, which takes the value 1 if a judge was a member of the communist-leaning Young Jurists League (YJL) in 1969 and 0 otherwise.² They argue that leftist judges are discriminated in moving up the pay scale, and thus expect that in a regression model of *Time2Sok*, the coefficient for *YJL* is estimated to be statistically significant positive. In addition, they consider the following factors as control variables.

Flunks is defined as the number of years between college graduation and entrance to the Legal Research and Training Institute, which is estimated from birth year and assumed to approximate the number of times a judge failed the entrance exam to the Legal Research and Training Institute.³ *Elite_College* takes the value 1 if a judge graduated from either the University of Tokyo or the University of Kyoto, two of the most prestigious national universities in Japan, and 0 if otherwise.⁴ *Ist_Tokyo* takes the value 1 if a judge started at the Tokyo District Court and 0 otherwise, accounting for the fact that the most promising are generally assigned to the Tokyo District Court.⁵ In order to control for unobservable differences among the cohorts, dummy variables indicating the year in which a judge finished his or her legal education are incorporated into estimation.⁶

Model 1 in Table 1 reports the OLS results in Ramseyer and Rasmusen (2003, p.41).⁷ What most notable from their perspective is the coefficient of 0.919 for *YJL*, whose statistical significance is 0.088. The estimate implies that judges who were the YJL members as of 1969, controlling for other factors, received their first DCJ assignment roughly a year later than their non-YJL peers. Although this is not the sole evidence they provide, it is certainly the most fundamental one to support their idea of conservative judges doing better in climbing the pay

² Ramseyer and Rasmusen use Shiso Undo Kenkyusho (1969), which reproduced the list of members from the YJL’s newsletter.

³ The two cases with no birth record are assigned with the mean of *Flunks* for the group as a whole. The calendar year is used to calculate *Flunks*.

⁴ In cases with no school record, judges are assumed to be non-elite college graduates.

⁵ A judge started at a summary court in the Tokyo area is not considered as a Tokyo starter.

⁶ Ramseyer and Rasmusen also examine another estimation model that includes variables identifying whether or not a judge is male and whether or not a judge started at a branch office, although we do not consider these and focus on their simpler model.

⁷ Following Ramseyer and Rasmusen, the estimates for the cohort dummies are not reported. The data used to estimate the model is available at

<http://pacioli.bus.indiana.edu/erasmuse/jbook/jbook.htm>.

scale. We believe that it is worth checking whether or not the estimated effect of *YJL* on *Time2Sok* has a systematic basis.

<Table 1 about here>

RECONSIDERATION

1. Data Cleaning

Given our goal of reconsidering *Time2Sok* as a duration time to reach the first DCJ post in a survival analysis setting, we compiled the judicial career data by ourselves. In the process of data compilation, we found that some of their data do not exactly match what are available in the data sources. To be concrete, mismatches are identified for 15 cases for *Time2Sok*, 4 cases for *YJL*, 1 case for *Elite_College*, and 4 cases for *Flunks*.⁸ Model 2 in Table 1 reports the OLS results using our data set with corrections. These estimates are not drastically different from what Ramseyer and Rasmusen reported. Their data and our corrected data are not substantially different, as indicated by Table 2 that reports the summary statistics of the variables used to estimate Models 1 and 2 in Table 1, except the cohort dummies. Nonetheless, the estimate of 0.818 for *YJL* is slightly smaller, and with the almost same standard error, considerably losing the statistical significance (0.131). Thus, even without reconsidering the adequacy of regressing *Time2Sok* in an OLS setting, it is clear that their argument of *YJL* members moving up the pay scale more slowly is not warranted and critically depends on whether or not their data set was compiled with care.

<Table 2 about here>

2. Survival Analysis

Duration and Right Censoring. Moreover, the analysis by Ramseyer and Rasmusen relies heavily on their assumption of *Time2Sok* and the adjustments made when they compiled the career data. Why had they excluded the “dropouts” from the career data, who quit or died early in their career without obtaining a DCJ appointment, and assumed even the “losers” who never obtained a DCJ appointment reaching their first DCJ appointment at their death or resignation? It is simply because without these adjustments, they cannot calculate *Time2Sok* for these “dropouts” and “losers.” However, these adjustments are problematic since the “dropouts” had a chance to receive a DCJ posting if they were not dead or continued to serve as a judge, and the exclusion of these “dropouts” ignores the time they spent as a judge to reach

⁸ See Masuyama (2005) for details. One of the mismatches for *Time2Sok* is the case that took 17 months to become a judge after the graduation from Legal Research and Training Institute, which should not be included in the data set if the criterion of Ramseyer and Rasmusen is strictly applied. It seems they used the year the judge hired as a starting year to calculate his *Time2Sok*.

DCJ. Similarly, quasi *Time2Sok* calculated for the “losers” is necessarily shorter than their true *Time2Sok* since the latter must include the additional time they needed to receive a DCJ posting if they were not dead or continued to serve as a judge. This problem is more serious for the judges who were assumed to obtain a DCJ appointment in 1997, due to the list of judges that only covers the job postings up to 1997. In other words, these data adjustments affect in a way to systematically exclude the time spent as a DCJ candidate and necessarily cause bias in estimating time to DCJ.

Survival analysis is a set of statistical techniques used to provide an answer to the question of how long it takes for an event of interest to occur.⁹ To be concrete, survival analysis asks how the duration spent in one state affects the probability that an individual of interest will make a transition to another state. The focus of survival analysis is placed on estimating the hazard rate, which is the probability that an event will occur at a particular point in time, given that it has not yet occurred. Not only survival analysis deals simultaneously with the questions of if something occurs and when it occurs, but also it provides an efficient solution for the issue of censoring, which becomes a problem whenever the scope of analysis is limited to a certain time period. In survival modeling, an event of interest is unobservable simply because it has not yet reached a point in time for the event to occur. In other words, we are unable to observe the event of interest that otherwise may be observed if the scope of analysis were somewhat longer or the subject remained to be under observation. Survival analysis treats these unobserved cases as “censored” to be dealt with in a manner compatible with the cases for which the event of interest is observed.

In order to conduct survival analysis of time to DCJ using the entire career data for the period between 1959 and 1968, we need to introduce a set of new variables. First, *Event* takes the value 1 if a judge received a DCJ posting and 0 otherwise. While Ramseyer and Rasmusen dropped the judges who held non-judicial postings before their first DCJ posting, we keep these “stars” in our data set (thus *Event*=1) and take into account the time they held non-judicial postings with the method we will explain later. Also, we keep the “dropouts” who never received a DCJ posting and quit or died before 20.41 years in our data set (*Event*=0).

Second, *Time* is defined as the years a judge needed to receive a DCJ appointment (or censored without reaching DCJ) after his or her graduation from the Legal Research and Training Institute. Note that both *Time* and *Time2Sok* measure the time a judge takes to reach the first DCJ post, although the former is one year greater than the latter so that the former is amenable to logarithmic transformation.

Let E , T and T^* denote *Event*, *Time*, and *Latent Time*, respectively. *Latent Time* is the years a judge would have needed to receive a DCJ appointment if not censored. We assume T^*

⁹ See Box-Steffensmeier and Jones (2004) for survival analysis in social science.

follows a log logistic distribution, $f(T^*)$. A scale parameter is re-parameterized by using covariates W and coefficients b : $\text{Median}(T^*) = \exp(bW)$. If we observe that a judge becomes DCJ ($E=1$), its contribution to likelihood is proportional to $f(T)$. If the duration of a judge remaining as a DCJ candidate is censored ($E=0$), its contribution to likelihood is proportional to the probability of survival:

$$S(T) = \int_T^{\infty} f(T^*) dT^* .$$

Parameters are estimated by maximizing total likelihood.

Left Truncation. In our data set, if a judge held a non-judicial post (i.e., officers in the Secretariat of the Supreme Court or the Ministry of Justice), the time to receive a DCJ posting is treated as censored since s/he is no longer at “risk” of reaching a DCJ post and his or her duration time is terminated without a transition to DCJ. When s/he returned to a judicial post, his or her time to DCJ resumed, although the resumed time to DCJ includes the time s/he was not at risk of reaching a DCJ post, which needs to be considered separately from the time s/he was at risk of obtaining a DCJ appointment. This is the issue addressed as “left truncation” in survival analysis, and we introduce a new variable, *Begin*, to account for the time not at risk of receiving a DCJ posting, which is defined as the years a judge took to resume a judicial posting. *Begin* increases along with the years s/he needed to resume a judicial posting, and needs to be equated with the years passed since his or her graduation from the Legal Research and Training Institute.

Ramseyer and Rasmusen dropped judges who did not join the court within a year of finishing their legal training (mostly those first became a prosecutor or attorney and later joined the court). We follow their strategy in this regard and additionally drop three judges whose birth years remain unknown and three other judges who died before 1969 and had no chance of becoming a YJL member. Consequently, our career data set consists of 717 judges, which generates 998 observations of time to DCJ, and the median, minimum, and maximum of *Time* for the enlarged data set is respectively 18, 2, and 38 years.¹⁰

When time to DCJ starts at T_0 , both survival probability and failure density are conditioned on the fact that an event has not happened by T_0 , whose survival probability is $S(T_0)$. Thus, these conditional survival probability and failure density at $T > T_0$ are $S(T)/S(T_0)$ and $f(T)/S(T_0)$, respectively.

¹⁰ Appendix to Masuyama (2005, 411-435) incorrectly lists the following judges as non-YJL members (ID in parentheses): Abe Akira (1102), Kuwata Katsutoshi (1127), Oguri Takao (1220), Yoneda Yasukuni (1285), Sawada Osamu (1334), Hirokawa Kouji (1574), and Inoue Jirou (1603).

Split Population. Ramseyer and Rasmusen implicitly assume that all judges eventually become DCJ, namely, they are at risk of reaching DCJ. This assumption is dubious, however. If leftist judges are severely discriminated, some of them are never on the promotion track. In split population modeling, the population of observations is divided into two groups, one that will experience an event of interest and one that will not. Let r be the probability for a judge i to be at risk of a transition or on the promotion track. Then, the probability to be censored at T is $(1-r) + rS(T)$. The probability to have an event of interest at T is $rf(T)$. Moreover, we can re-parameterize r by using some covariates Z and coefficients a : $r = (1+\exp(aZ))^{-1}$. We use R to analyze the judicial career data and write our own code to deal with the issue of left truncation in split population model.¹¹

3. Time to Higher Positions

Once we regard time to DCJ as a duration time with or without being censored, we can also analyze time to other positions in a survival analysis setting. The reason why Ramseyer and Rasmusen examine time to DCJ is that DCJ is a moderately prestigious administrative status, which can be used as a proxy for pay. It is true that the Secretariat promotes some judges earlier to DCJ than others. But, it should also be the case for other higher positions. As Ramseyer and Rasmusen (2003, p. 24) themselves admit, High Court presidents and District Court chief judges are *more* prestigious and with *higher* pay. Discrimination, if any, would be clearer for these posts than for DCJ. Since not so many judges are promoted to these posts, it is hardly an adequate question to deal with their pay in an OLS setting.

Moreover, some judges skip to High Court without becoming DCJ at a District Court. On the other hand, some branch chief judges and non-judicial posts may be considered as an equivalent to DCJ. Thus, DCJ may not be an ideal post to analyze pay discrimination.

In contrast, we have no need to truncate the period during which judges held non-judicial posts because they are continuously at risk of reaching such a higher position (i.e., a High Court president and a District Court chief judge). Hence, every judge has one duration observation to each post, and thus we distinguish these non-left-truncated data (N=717) from the previous left-truncated data (N=998). For these positions, dependent variables *Time* and *Event* are defined as in the case of DCJ except data collection period is up to 2003.

4. Causal Inference for Survival Analysis

Model Dependence. Even if we sincerely try to specify the model, we are unlikely to do it perfectly. Failure time may not follow a log logistic distribution and the split population

¹¹ Box-Steffensmeir, Radcriff and Bartels (2006) offers a STATA program to run split population model, although their program does not consider left truncation.

assumption may not be valid, though we can not be sure. In this sense, we will have model dependent results from application of any parametric model to raw data (As for strict causal inference in political science, see Diamond and Sekhon (2004), Ho et al. (2004) and King and Zeng (2006)).

If we take causal inference seriously, the effect of YJL on latent promotion time T^* should be a difference between T_i^* of judge i who would belong to YJL and T_i^* of the *same* judge i who would not belong to YJL. Let X be a dummy variable of YJL (which is called treatment in the literature of causal inference), M be a vector of all measured covariates. That is, realized causal effect for judge i is

$$\Delta T_i^*(M_i) = T_i^*(X_i=1, M_i) - T_i^*(X_i=0, M_i).$$

If we assume that no omitted variables are associated with X_i and affect T_i^* , this difference is solely caused by a difference of X_i 's value. Nonetheless, "the fundamental problem of causal inference" (Holland 1986) arises since we cannot observe both $T_i^*(X_i=1)$ and $T_i^*(X_i=0)$ but at most either of the two. If $X_i=1$ and an event of interest occurs ($E_i=1$), we observe $T_i^*(X_i=1, M_i) = T_i$ but not $T_i^*(X_i=0, M_i)$. If $X_i=0$ and $E_i=1$, we observe $T_i^*(X_i=0, M_i) = T_i$ but not $T_i^*(X_i=1, M_i)$. Otherwise, we observe neither of the two.

However, we are interested not in the realized effect on value T_i^* of judge i but in a potential effect on a random variable T^* of judge population. One of the quantities of interest for most scholars is called "average treatment effect" in the causal inference literature, and is expressed as

$$\frac{1}{n} \sum_{i=1}^n E(T_i^* | X_i=1, M_i) - E(T_i^* | X_i=0, M_i).$$

Suppose that we assume $E(T^*|X, M)$ has a functional form $g(X, M)$ such as a linear function or logistic link. After we estimate parameters in $g(\cdot)$ by using a conventional survival model, we can estimate an average treatment effect by calculating

$$\frac{1}{n} \sum_{i=1}^n g(X_i=1, M_i) - g(X_i=0, M_i).$$

A problem is that the model misspecification arises if $E(T^*|X, M)$ is not equal to $g(X, M)$ and thus the estimate differs from the estimand.

Matching. One way to avoid this problem is matching. Given the assumption of no omitted variable, T_i^* of a hypothetical non-YJL judge i (but, in fact, $X_i=1$ is observed) should be equal to that of an observed non-YJL judge j ($X_j=0$) whose covariates M have the same values as those of judge i ,

$$T_i^*(X_i=0, M_i) = T_j^*(X_j=0, M_j=M_i).$$

We can estimate unobserved $T_i^*(X_i=0, M_i)$ by $T_j^*(X_j=0, M_j=M_i) = T_j$ if the latter is observed, namely, $E_j=1$. As often the case with survival analysis, we assume censoring is

non-informative, namely, E_i is independent of T_i^* . Accordingly, even if we drop such a matched pair (not just an observation) where at least one judge's duration is censored and if we calculate an average treatment effect, our estimates will not be biased (but inefficient). Note that, if data were not matched one, dropping censored cases would induce bias (underestimate potential duration of censored observations). In a matched data, when we drop a censored observation, we also drop the corresponding observation. Therefore, no bias occurs. Let $j=m(i)$ so that $M_i=M_j$ and $X_i=1-X_j$. We propose to estimate an average treatment effect on duration time by

$$\frac{1}{n(+)} \sum_{i=1}^{n(+)} [X_i \{T_i(X_i=1, M_i, E_i=1) - T_i(X_{m(i)}=0, M_{m(i)}=M_i, E_{m(i)}=1)\} + (1-X_i) \{T_i(X_{m(i)}=1, M_{m(i)}=M_i, E_{m(i)}=1) - T_i(X_i=0, M_i, E_i=1)\}],$$

where $n(+)$ is the number of observations where $E_i=1$ and $E_{m(i)}=1$.

Matching also reduces model dependence in parametric modeling. Ho et al. (2004) recommend that we use non parametric matching as “preprocessing,” namely, a tool to remove bias of correlated control variables, before we perform any parametric analysis. We still rely on a functional form $g(\cdot)$ and survival model specification such as a log logistic distribution for failure time and the split population assumption, though we need not worry about which independent variables are included and whether they should be squared or interacted with other variables. Even if matching fails to balance covariates completely, parametric models control covariates (as long as they are correctly specified).

It is, however, impossible to match every YJL judge with such a non-YJL judge who has the exactly same covariates (exact matching). As a second best method, we will match every observation to an observation which is as similar as possible with replacement (nearest neighbor matching). Difficult part is how to detect such pairs and this is why various matching methods are proposed. This paper employs a genetic matching method by Diamond and Sekhon (2004). We will assess the performance of matching by checking how well distributions of covariates (as well as a joint distribution of them) are balanced between YJL judges and the others.

REANALYSIS

Matching and Non Parametric Analysis

When we match observations in left-truncated data, we match on *Flunks*, *Elite_College*, *Begin*, *Strata* (the number of previous duration observations of the same judge),¹²

¹² For example, suppose that a judge returns to a judicial post after holding a non-judicial post. In this case, time to non-judicial post is the first duration for this judge, while the resumed

their interactions and squares (except squared *Elite_College*), ten cohort dummies, interactions between cohort dummies and the other four variables. 1997 pairs are matched on. In the case of non-left-truncated data, we match on the above mentioned variables except *Begin*, *Strata*, their squares and interactions. We do not match on *Ist_Tokyo* because it is affected by *YJL* (post treatment variable). 1664 pairs are matched on.¹³

Balance of covariates between *YJL* judges and non *YJL* judges is almost achieved in both datasets. For example, seven cohort dummies and five interactions between cohort dummies and *Elite_College* are perfectly balanced in the left-truncated data. Even for the worst balanced variable, *Flunks*, before-matching mean values are 3.49 and 4.12 for *YJL* judges and non *YJL* judges respectively, while those for after-matching are 3.65 and 3.91; the difference becomes narrower.

In order to summarize the balance of covariate distributions, we compare the distributions of propensity scores between *YJL* judges and non *YJL* judges before and after matching. Propensity scores are predicted values of logistic analysis where we regress *YJL* on all matched-on covariates, implying the probability that the observation is an *YJL* judge. When balance is achieved and a distribution of covariates is independent of *YJL*, the distributions of propensity scores should not differ between the two judge groups. According to Figure 1 (left-truncated data) and Figure 2 (non-left-truncated data), the distributions of propensity scores are considerably different before matching, while they are similar after matching. Clearly, covariates are better balanced after matching in both the left-truncated data and the non-left-truncated data.

<Figures 1 and 2 about here>

Figure 3 depicts life tables. According to the upper panel, survival curves are indistinguishable between the two groups up to 25th year, while they differ after that. It is difficult to decide whether *YJL* affects *Time* to DCJ. The lower panel tells the two lines are similar and *YJL* does not matter for *Time* to District Court chief judge. Since only 30 of 717 judges became High Court president in our data set, its survival curves are almost flat and identical between the two groups. We omit the figure for High Court president.

<Figure 3 about here>

Figures 4 and 5 respectively show the distributions of treatment effect estimates of *YJL* on DCJ and District Court chief judge. Average treatment effects are 0.253 and -0.093 (shown as vertical lines), but their standard deviances are 5.94 and 2.76. We cannot say that *YJL* has any effect on *Time* to either of these posts. Since only one matched pair has a set of observed

duration is the second one.

¹³ We use a statistical environment R for estimation. For matching, we use Matching library. Options are as follows; estimand is ATE; boot is 1000 following recommendation for publication standard.

time to High Court president for YJL and non YJL judges, we do not estimate average treatment effects on this post.

<Figures 4 and 5 about here>

Survival Analysis

Tables 3, 4 and 5 respectively report the results of survival analysis of *Time* to DCJ, District Court chief judge and High Court president. We calculate standard errors by non-parametric bootstrapping. We resample matched pairs, not observations, with replacement in order to retain a balanced joint distribution of covariates between the two groups. We repeat this resampling 1000 times, calculate standard deviances of 1000 sets of parameter estimates in the survival analysis, and multiply them by $1000/(1000-1)$, which will be our estimates of standard errors. The significance level is based on quantile values without assuming any distribution of estimates. Note that standard errors are not robust against deviant values and coefficients are sometimes significant despite the relatively larger standard errors.

<Tables 3, 4 and 5 about here>

For the covariates in both duration and split population models, we use the same independent variables introduced by Ramseyer and Rasmusen (2003). Among the coefficients of *YJL*, only that of split population part of DCJ analysis is significantly different from zero at 1 % level. This means *YJL* significantly decrease the probability that a judge is at risk of becoming DCJ. Examining the life table in Figure 3 again, some YJL judges have no chance to reach DCJ position but, once they are on the track to DCJ, they are not discriminated in terms of timing to DCJ. This result does not contradict Figure 4, which corresponds to the duration part of Table 3. When it comes to the other two posts, *YJL* does not matter.

To briefly discuss the estimates for other independent variables, *Ist_Tokyo* fastens *Time* to DCJ and increases the probability to be on the track to District Court chief judge and High Court president. *Flunks* decreases chance to be in the population of no DCJ risk. Since *Flunks* is completely correlated with age when a judge starts his/her career, this result may imply that, once relatively older judges are in the DCJ population, they reach DCJ earlier due to their age. *Elite_College* makes a judge more likely to be on the track to High Court president.

CONCLUSION

Is the Japanese judicial system independent of political control? Ramseyer and Rasmusen (2003) studied how long it takes for a judge to reach a DCJ post and argued that judges with leftist preferences do worse in their careers. However, their analysis is problematic since 1) it is based on inaccurate data, 2) takes no account of the issues of censoring

and time dependence, and 3) critically depends on their model (mis-)specification. In this paper, we proposed a matching method to reduce model dependence, calculated an average treatment effect of being a member of YJL on the timing to reach DCJ, and employed survival analysis to deal with the issues of right censoring, left truncation and split population. Our analysis shows that a judge who once belonged to YJL is somewhat less likely to be on the track to DCJ, but does not take longer to reach DCJ once s/he is on the promotion track. Our findings are confirmed by the absence of average treatment effect for YJL and non YJL judges, even considered other prestigious judicial posts such as District Court chief judges and High Court presidents.

Our method is not limited to an application to Japanese judges. If the problem at hand is the occurrence of an event such as an employee obtaining a certain rank, we must properly deal with the issues of right censoring. We have not only applied survival analysis to the data set of judicial careers, but also taken into account the possibility of not all judges experiencing the event of interest. Political scientists become more and more sensitive to model dependence. We have shown that matching helps the results of our survival analysis to be robust against model misspecification, and proposed a way to estimate an average treatment effect on the timing of a judge receiving a certain appointment. Although much remains to be done, we believe that our analysis provides a good example of causal inference for the students of both judicial politics and political methodology.

Table 1. OLS Results of Time to DCJ

	Model 1			Model 2		
	Coef.	S.E.	Sig.	Coef.	S.E.	Sig.
<i>YJL</i>	0.919	0.538	0.088	0.818	0.540	0.130
<i>Ist_Tokyo</i>	-1.383	0.804	0.086	-1.335	0.793	0.093
<i>Flunks</i>	0.014	0.071	0.849	0.019	0.070	0.790
<i>Elite_College</i>	0.086	0.516	0.867	0.086	0.522	0.870
<i>Constant</i>	22.905	0.785	0.000	22.887	0.787	0.000

Note: N=501. The coefficients for cohort dummies are omitted. Significance is based on a two-tailed test. For both models, adjusted R² is 0.09.

Table 2 Summary Statistics of OLS Models

	Ramseyer and Rasmusen data				Fukumoto and Masuyama data			
	Mean	Std	Min	Max	Mean	Std	Min	Max
<i>Time2Sok</i>	21.711	5.317	11	37	21.681	5.364	11	37
<i>YJL</i>	0.281	0.450	0	1	0.289	0.454	0	1
<i>Ist_Tokyo</i>	0.104	0.305	0	1	0.108	0.310	0	1
<i>Flunks</i>	4.828	3.441	0	18	4.880	3.505	0	18
<i>Elite_College</i>	0.317	0.466	0	1	0.319	0.467	0	1

Note: N=501.

Table 3. Survival Analysis of Time to DCJ

	Duration			Split	
	B	SE		B	SE
<i>YJL</i>	0.012	0.014		-0.872	0.869 **
<i>Ist_Tokyo</i>	-0.086	0.022	**	0.488	4.734
<i>Flunks</i>	-0.009	0.003	**	-0.273	0.110 **
<i>Elite_College</i>	-0.002	0.013		-0.366	0.355
<i>Log(Shape)</i>	-2.147	0.023	**		
<i>Constants</i>	3.210	0.018	**	4.239	1.752 **

Note: N=3994. Log likelihood is -6003.64.

* p < 0.1 ** p < 0.05

Standard errors and significance are based on bootstrap.

Log logistic model is used.

The coefficients for cohort dummies are omitted.

Table 4. Survival Analysis of Time to District Court Chief Judge

	Duration		Split			
	B	SE	B	SE		
<i>YJL</i>	0.006	0.019	-0.208	10.649		
<i>Ist_Tokyo</i>	-0.034	0.02	0.662	9.484	*	
<i>Flunks</i>	0.009	0.008	-0.098	3.825		
<i>Elite_College</i>	-0.022	0.021	0.393	8.486		
<i>Log(Shape)</i>	-2.969	0.297	**			
<i>Constants</i>	3.523	0.017	**	3.211	13.628	**

Note: N=3328. Log likelihood is -680.8837.

* $p < 0.1$ ** $p < 0.05$

Standard errors and significance are based on bootstrap.

Log logistic model is used.

The coefficients for cohort dummies are omitted.

Table 5. Survival Analysis of Time to High Court President

	Duration		Split		
	B	SE	B	SE	
<i>YJL</i>	-0.031	0.028	-3.112	30.958	
<i>Ist_Tokyo</i>	-0.026	0.072	34.146	55.377	**
<i>Flunks</i>	-0.003	0.011	-0.601	14.423	
<i>Elite_College</i>	-0.025	0.881	20.765	37.243	*
<i>Log(Shape)</i>	-4.347	0.401			**
<i>Constants</i>	3.911	1.197	-44.453	47.028	**

Note: N=3328. Log likelihood is -59.06214.

* $p < 0.1$ ** $p < 0.05$

Standard errors and significance are based on bootstrap.

Log logistic model is used.

The coefficients for cohort dummies are omitted.

Figure 1. Distributions of Propensity Scores: Left-Truncated Data

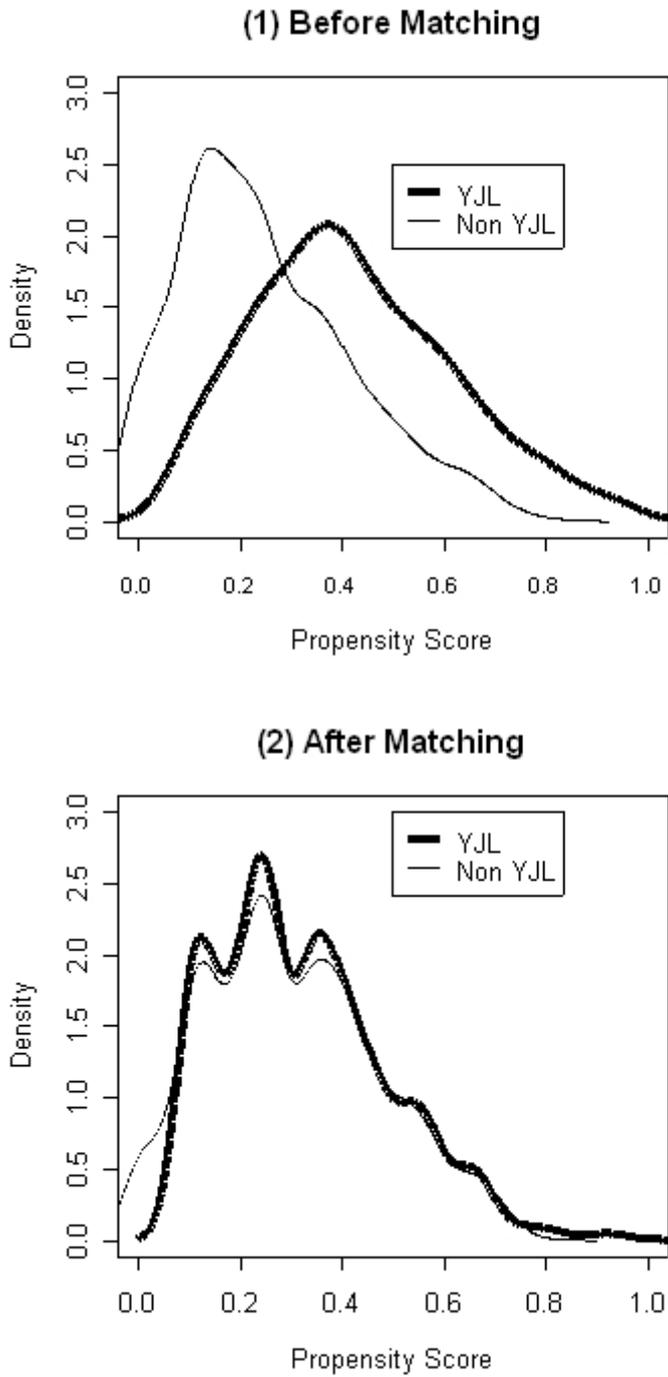


Figure 2. Distributions of Propensity Scores: Non-Left-Truncated Data

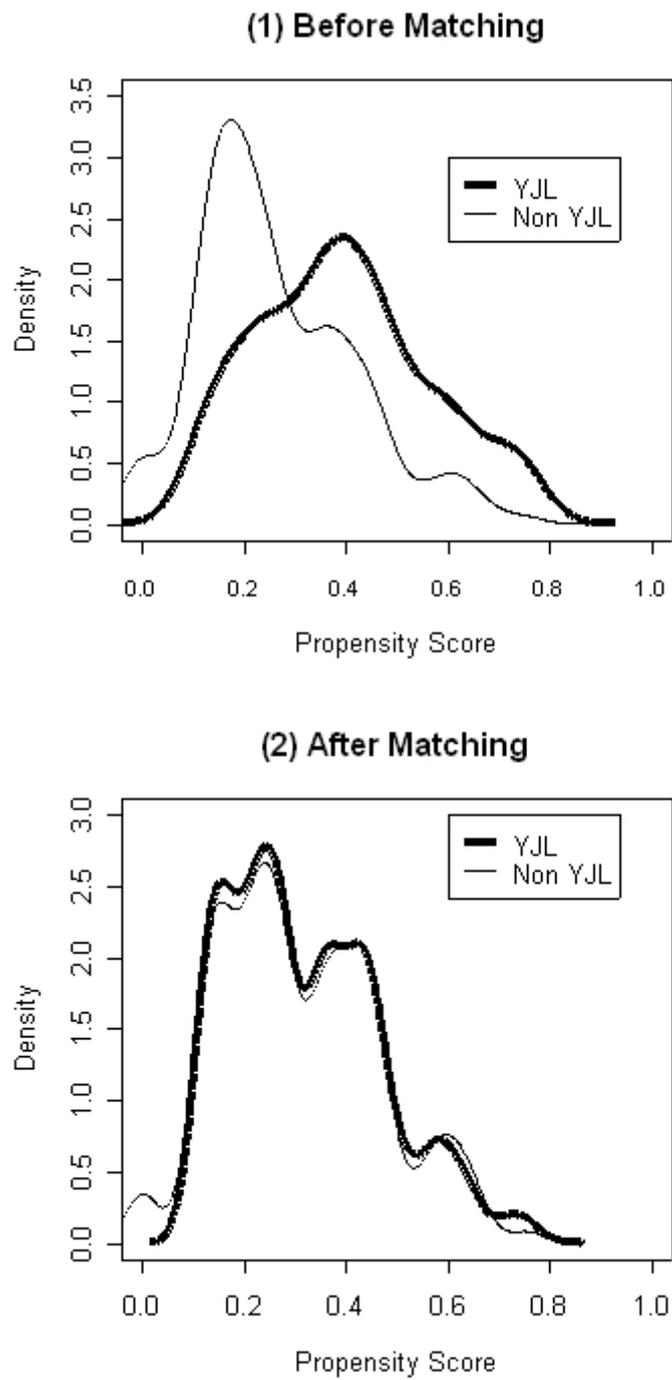


Figure 3. Comparison of Life Tables between the Two Judge Groups

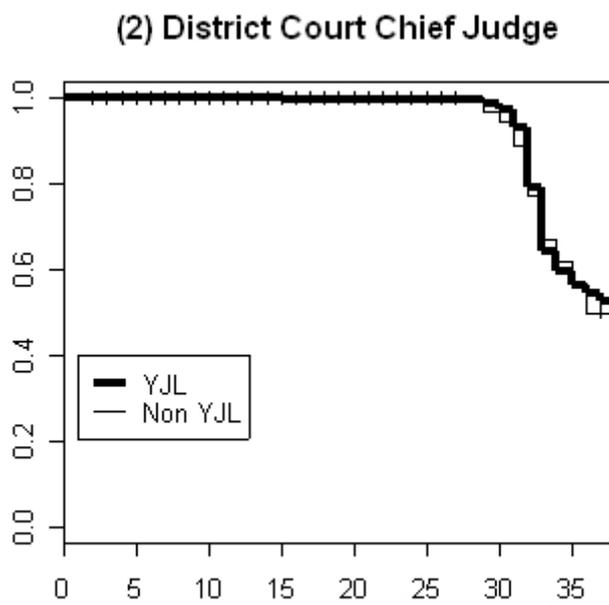
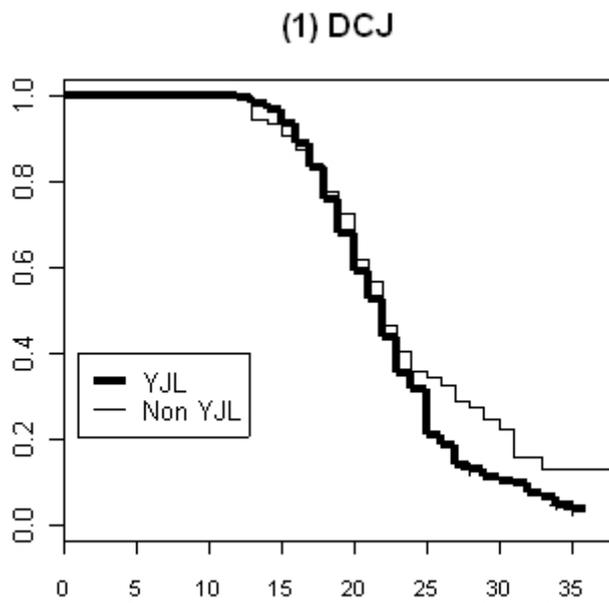


Figure 4. Effect of YJL on DCJ

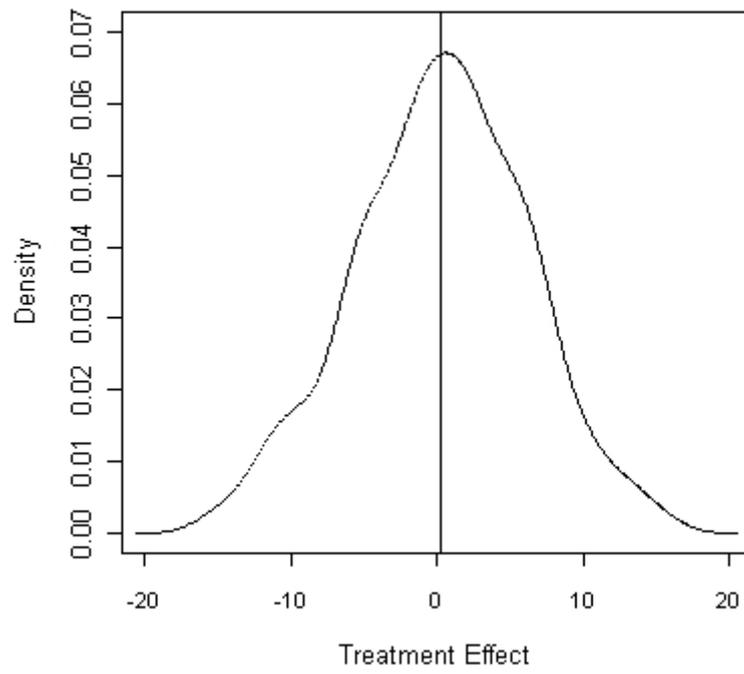
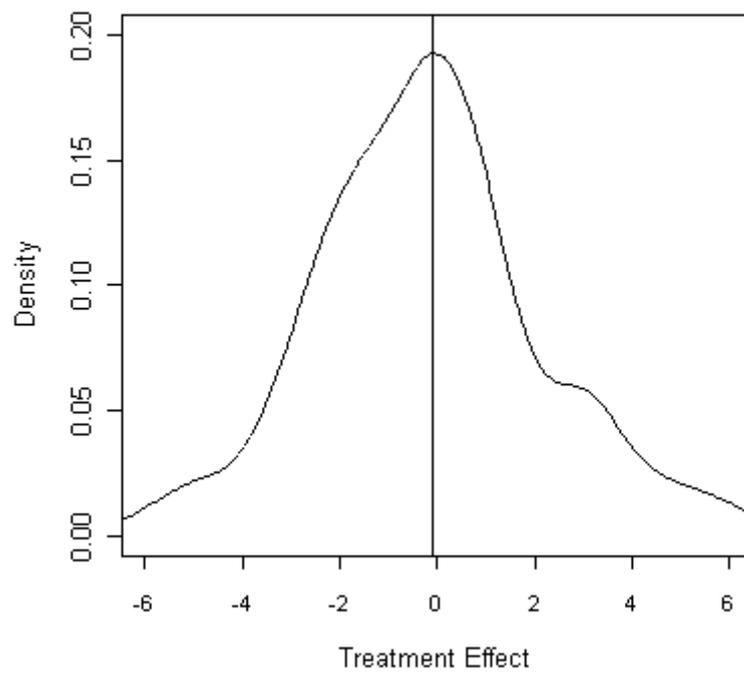


Figure 5. Effect of YJL on District Court Chief Judg



REFERENCE

- Box-Steffensmeier, Janet and Bradford Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.
- Box-Steffensmeier, Janet M., Peter M. Radcliffe, and Brandon L. Bartels. 2006. "The Incidence and Timing of PAC Contributions to Incumbent U.S. House Members, 1993-94." *Legislative Studies Quarterly* 30: 549-79.
- Diamond, Alexis, and Jasjeet S. Sekhon. 2004. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." Harvard University. Typescript.
- Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth Stuart. 2004. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." Harvard University. Typescript.
- Holland, Paul W. 1986. "Statistics and Casual Inference." *Journal of American Statistical Association* 81: 945-960.
- King, Gary, and Langche Zeng. 2006. "When Can History be Our Guide? The Pitfalls of Counterfactual Inference." *International Studies Quarterly* forthcoming.
- Masuyama, Mikitaka. 2005. "Saibankan Kyaria no Seizon Bunseki." *Seikei Hogaku*. 61:411-456.
- Nihon Minshu Horituka Kyokai Shiho Seido Iinkai. 1998. *Zensaibankan Keireki Soran: 3rd ed.* Tokyo: Kojinsha.
- Ramseyer, Mark and Eric Rasmusen. 1997. "Judicial Independence in a Civil Law Regime: The Evidence from Japan." *Journal of Law, Economics and Organization* 13: 259-286.
- Ramseyer, Mark and Eric Rasmusen. 2001. "Why Are Japanese Judges So Conservative in Politically Charged Cases?" *American Political Science Review* 95: 331-344.
- Ramseyer, Mark and Eric Rasmusen. 2003. *Measuring Judicial Independence: The Political Economy of Judging in Japan*. Chicago: University of Chicago Press.
- Shiso Undo Kenkyusho. 1969. *Osorubeki Saiban*. Tokyo: Zenboshu.