GAAR in Action: Judicial Attributes and Transactions Types

Thaddeus Hwong, for a collaborative project with Tim Edgar and Jinyan Li[*]

Here is a rough summary of the latest set of probit regression results prepared by Thaddeus Hwong for the symposium. The data used are the cumulation of the work of Osgoode Hall Law School student research assistants made possible by the generous funding support from the Canadian Tax Foundation. The data they coded will be checked and updated again for the development of the full paper.

Tim Edgar once modestly said he had only a few epiphanies in his life, and his idea about the classification of GAAR cases by transaction types was one of them. As he wrote over a decade ago, his typology classifies GAAR cases into those featuring transactions that attempt to create a tax attribute, those featuring transactions that attempt to transfer a tax attribute and those featuring transactions that involve the substitution of a lower-taxed transactional form for a higher-taxed transactional form.[1] In 2011, he led our first application to seek funding from the Canadian Tax Foundation to run empirical tests of his typology in the context of judicial decision making in a cross-country setting. His view that there's room to make a contribution in advancing knowledge with the project looms large in what he subsequently wrote in 2014 as editor of *Canadian Tax Journal*: "It is probably accurate to suggest that GAAR has altered the contours of tax planning in Canada. That altered landscape may be attributable in part to the fact that the jurisprudence considering the interpretation and application of this provision does not readily yield an especially clear road map, at least in the sense of providing certainty and predictability of result."[2] With the generous funding support from the Canadian Tax Foundation, here is a set of the latest empirical results.

1   Modelling Effects of Transaction Types on Judicial Decision Making in GAAR Cases

To empirically explore what the three transaction types could contribute to predicting how judges would decide GAAR cases, data were collected concerning GAAR cases decided in Canada, Australia and New Zealand as well as the judges who decided them for the development of a unique GAAR case-judicial dataset in which the data about the cases and the data about the judges are linked.

The latest version of the dataset contains data about the cases and the judges who decided them from 72 Canadian cases in 1997-2017 (31 held for the government), 78 Australian cases in 1988-2013 (44 held for the government, and 36 New Zealand cases in 1982-2014 (30 held for the government).

A research approach of an exploratory nature anchored on probit regression analysis was adopted. The judges' decisions were coded as either votes cast for the government or votes not cast for the government. Votes cast by the judges in the cases were used as the dependent variable. Five categorical variables and one continuous variable were used as the independent variables.

In line with the standard probit regression setup for categorical variables, one category of each categorical variable was designated as the base category used as the basis by other categories of the same variable to compare against. The categorical independent variables were set up as follows:

- the transaction type in each case (the creation of tax attributes as the base category in comparison with the trading of tax attributes and the substitution of tax attributes);
- the jurisdiction in which each case was decided (Canada as the base category in comparison with Australia and New Zealand);
- a proxy of the political climate of the country in which each case was decided (when conservative governments were in power as the base category in comparison with when liberal governments were in power);
- a proxy of the political leanings of the judges (judges appointed by conservative governments as the base category in comparison with judges appointed by liberal governments); and
- the gender of each judge (male judges as the base category in comparison with female judges; given the availability of raw data only the male/female differential was coded).

The continuous independent variable was the judicial experience of each judge as represented by the number of years of the judge on the bench when the judge decided a case.

2   Describing the Case-Judicial Data By Variables Used In Probit Regression Analysis

In describing the data used for the probit regression modelling, the presentation here is designed only to share a snapshot of the data used.[3] To streamline the presentation, cases featuring the creation of tax attributes are referred as `Creation` cases, cases featuring the trading of tax attributes are referred to as `Trading` cases, while cases featuring the substitution of tax attributes are referred to as `Substitution` cases.

The Canadian cases were decided by 51 judges who cast 102 votes, of which 53 were cast for the government while 49 were not cast for the government. Among the 102 votes, 67 (65.7%) were cast in `Creation` cases, 10 (9.8%) were cast in `Trading` cases, while 25 (24.5%) were cast in `Substitution` cases. Most votes were cast in `Creation` cases. In fact, the sum of the number of votes cast in `Trading` cases and `Substitution` cases is roughly about half of the number of votes cast in `Creation` cases.

The Australian cases were decided by 67 judges who cast 172 votes, of which 96 were cast for the government while 76 were not cast for the government. Among the 172 votes, 55 (32%) were cast in `Creation` cases, 17 (10%) were cast in `Trading` cases, while 100 (58%) were cast in `Substitution` cases. Votes cast in `Substitution` cases outnumbered cases featuring the other two types combined.

The New Zealand cases were decided by 50 judges who cast 87 votes, of which 72 were cast for the government while 15 were not cast for the government. Among the 87 votes, 43 (49%) were cast in `Creation` cases, 5 (6%) were cast in `Trading` cases, while 39 (45%) were cast in `Substitution` cases. Votes cast in `Creation` cases and `Substitution` cases together account for almost all the votes cast. Even the number of judges of Canada and New Zealand covered in this exercise is similar, their disposition as reflected in the votes they cast is very different.

---

[3] Additional details of the data used would be deposited in the form of appendices at www.yorku.ca/thwong/types.

In sum, votes cast in **Trading** cases account for the smallest portion of total votes in each country. Votes cast in **Creation** cases account for a bigger portion of total votes than those cast in **Substitution** cases in Canada but vice versa in Australia. Votes cast in **Creation** and **Substitution** cases account for similar portions of total votes in New Zealand. Table 1 presents the votes of the three case types of each country by whether they were cast for the government or not.

Table 1: Judicial votes by outcomes

Judicial votes for government or not		Creation	(%)	Trading	(%)	Substitution	(%)	Total	(%)
Canada	**For**	42	63	5	50	6	24	53	52
	Not	25	37	5	50	19	76	49	48
	Total	67		10		25		102	
Australia	**For**	41	75	8	47	47	47	96	56
	Not	14	25	9	53	53	53	76	44
	Total	55		17		100		172	
New Zealand	**For**	36	84	1	20	35	90	72	83
	Not	7	16	4	80	4	10	15	17
	Total	43		5		39		87	

Canada. The only reason why the number of votes cast for the government and the number of votes not cast for the government are close is because the wider 42-to-25 margin of the two categories in **Creation** cases offsets the wide 19-to-6 margin of votes of the corresponding categories in **Substitution** cases.

Australia. The similarity between the Canadian cases and Australian cases is that votes cast for the government versus votes not cast for the government are sort of like mirror images in **Creation** cases and **Substitution** cases. The difference is that the spread in **Creation** cases and the spread in **Substitution** cases are narrower in Australian cases, and hence the wider spread in the total votes in Australia compared to the total votes in Canada.

New Zealand. New Zealand cases appear to be different from Canadian cases and Australian cases as a one-sided pattern emerges – most votes were cast for the government. Specifically, most votes cast in **Creation** cases and **Substitution** cases were cast for the government, and together they account for most of the votes cast.

In sum, most of the votes cast in **Creation** cases were cast for the government in all three countries. The impression of votes cast in **Substitution** cases is mixed – for the government in New Zealand by a wide margin, not for the government in Canada, not for the government in Australia but by a narrower margin than that in Canada. The impression of votes cast in **Trading** cases – as in all through this section – is less clear due to the small number of votes cast in each country. Table 2 presents the votes of the three case types of each country by whether they were cast when conservative governments or liberal governments were in power.

Table 2: Judicial votes by governments in power when cases were decided

Judicial votes cast when liberal/conservative governments in power		Creation	(%)	Trading	(%)	Substitution	(%)	Total	(%)
Canada	Lib	19	28	4	40	10	40	33	32
	Con	48	72	6	60	15	60	69	68
	Total	67		10		25		102	
Australia	Lib	22	40	0	0	53	53	75	44
	Con	33	60	17	100	47	47	97	56
	Total	55		17		100		172	
New Zealand	Lib	24	56	4	80	4	10	32	37
	Con	19	44	1	20	35	90	55	63
	Total	43		5		39		87	

Canada. Most votes were cast when conservative governments were in power, mainly because more than half of the votes cast in **Creation** cases were cast when conservative governments were in power.

Australia. The similarity between Australia and Canada is that more votes cast in **Creation** cases were cast when conservative governments were in power. But the spread between votes cast when conservative governments were in power and votes cast when liberal governments were in power is narrower in Australian **Creation** cases, and hence the narrower margin in the total number of votes in Australia. The difference between the two countries is that more votes were cast in the Australian **Substitution** cases when liberal governments were in power.

New Zealand. Again, one-sided patterns emerge. For votes cast when conservative governments were in power, the sum of the votes cast in **Creation** cases and **Substitution** cases account for all the votes cast except one vote cast in a **Trading** case. For votes cast when liberal governments were in power, votes cast in **Creation** cases alone account for almost all the votes.

In sum, more votes in each of the three countries were cast when conservative governments were in power. Any discernible cross-country pattern related to transaction types like the propensity to vote for the government in **Creation** cases prevailing in all three countries as depicted above in Table 1 is not detected. Table 3 presents the votes of the three case types of each country by whether they were cast by judges appointed by conservative governments or liberal governments.

Table 3: Judicial votes by governments that appointed the judges

Votes by judges appointed by liberal/conservative governments		Creation	(%)	Trading	(%)	Substitution	(%)	Total	(%)
Canada	**Lib**	32	48	4	40	11	44	47	46
	Con	35	52	6	60	14	56	55	54
	Total	67		10		25		102	
Australia	**Lib**	38	69	13	76	62	62	113	66
	Con	17	31	4	24	38	38	59	34
	Total	55		17		100		172	
New Zealand	**Lib**	22	51	3	60	14	36	39	45
	Con	21	49	2	40	25	64	48	55
	Total	43		5		39		87	

Canada. More votes were cast by judges appointed by conservative governments, but the spread between votes cast by those judges and votes cast by judges appointed by liberal governments is narrow in each of the three types of cases by transactions in terms of the number of votes.

Australia. Most votes were cast by judges appointed by liberal governments in all three types of cases. The spread between votes cast by judges appointed by liberal governments and those appointed by conservative governments is wider than that in Canadian cases.

New Zealand. The spread between votes cast by judges appointed by conservative governments and those appointed by liberal governments in **Substitution** cases is the reason why the total number of votes cast by the conservative-government appointees outnumbered the liberal-government appointees.

In sum, it's hard to say there's any distinct and consistent pattern related to transaction types emerged in the descriptive statistics here. Table 4 presents the votes of the three case types of each country by whether they were cast by male or female judges.

Table 4: Judicial votes by gender of judges

Votes cast by female/male judges		Creation	(%)	Trading	(%)	Substitution	(%)	Total	(%)
Canada	**Female**	18	27	3	30	4	16	25	25
	Male	49	73	7	70	21	84	77	75
	Total	67		10		25		102	
Australia	**Female**	5	9	0	0	17	17	22	13
	Male	50	91	17	100	83	83	150	87
	Total	55		17		100		172	
New Zealand	**Female**	4	9	1	20	4	10	9	10
	Male	39	91	4	80	35	90	78	90
	Total	43		5		39		87	

In Canadian cases, most votes were cast by male judges. In fact, male judges cast over 75% of the votes. Such male dominance rose to over 87% of votes in Australia and even 90% of the votes in New Zealand.

It's not an overstatement to say that no gender diversity among judges who decided GAAR cases is detected in all three countries.

In terms of the number of years of judges on the bench when a case was decided, the Canadian mean was over 10 years, with a high of 35 years; the Australian mean was 10 years, with a high of 27 years; and the New Zealand mean was 15 years, with a high of 31 years. As the median for each country is close to the mean respectively (Canada, 10 years; Australia, nine years; New Zealand, 15 years), the longevity of a small number of judges on the bench appears not to distract from the probit regression modelling as outliers that exert outsized influences.

3   Exploring Relationships Among Case-Judicial Data in Probit Regressions

Implementing the probit regression analysis, three models clustering on individual judges were run. The clustering took into consideration of the fact that a judge who cast more than one vote was the same individual who cast those votes. With the clustering, robust standard errors were used in the models. In all three models, the dependent variable was the votes cast by the judges in the cases. The independent variables were included in the three models in the following configurations:

Model 1. The only independent variable was the transaction types. Model 1 is the basic case attributes model on transaction types with no covariates.

Model 2. The independent variables representing the following two possible jurisdictional influences of judicial decision making – the jurisdictions in which cases were decided and the political climate of the countries in which cases was decided – were added to complement the independent variable of transaction types included in Model 1. Model 2 is the environmental model on transaction types with jurisdictions and their political climate as covariates.

Model 3. The independent variables representing the following three possible influences of the personal attributes of judges in judicial decision making – political leanings of the judges, their years on the bench when cases were decided and their gender – were added to complement the independent variables used in Model 2. Model 3 is the full model on transaction types with the independent variables about the jurisdictions plus the independent variables about the judges as covariates.

The above sequencing of the inclusion of the independent variables in the three models was designed to reveal any change in the effects of the independent variables given changes of the model configuration. Table 5 shows the results of the three models. From left to right, the columns are the categories of the independent variables that were compared against the corresponding base category, the effects of Model 1, those of Model 2 and those of Model 3. The number of observations used in each model is shown at the bottom of each column for each model. For each independent variable, two numbers are shown in two rows for each model. The top row shows the coefficient of the independent variable for a particular model. In probit regressions the coefficient is expressed in the change in z-score representing the change of the dependent variable given one unit of change of the independent variable. The statistically significant coefficients are marked with at least one asterisk, corresponding to one of the three levels of statistical significance listed at the bottom of the table. The bottom row shows the standard errors – the intuition is that the smaller the number is, the more precise the estimate will be.

Table 5: Probit Regression Results of the Three Models

| Judicial Votes	Model 1	Model 2	Model 3
Trading	−0.744**	−0.821**	−0.748**
	[0.261]	[0.275]	[0.282]
Substitution	−0.495***	−0.590***	−0.584***
	[0.132]	[0.146]	[0.147]
Australia		0.356*	0.319*
		[0.162]	[0.163]
New Zealand		1.057***	0.919***
		[0.193]	[0.198]
(a) Liberal governments in power when case were decided		−0.457**	−0.747***
		[0.142]	[0.224]
(b) Judges appointed by liberal governments			−0.163
			[0.172]
Interaction between (a) and (b)			0.642*
			[0.295]
Judges' years on the bench			0.0241*
			[0.0110]
Female Judges			−0.106
			[0.166]
Number of observations	**361**	**361**	**361**

Standard errors in brackets
* p<0.05, ** p<0.01, *** p<0.001

Six observations can be derived from the probit regression results. First, all except two independent variables – the variable representing whether judges were appointed by liberal or conservative governments and the variable representing the gender of the judges – were statistically significant. Second, the chance for judges to vote for the government in **Trading** cases and **Substitution** cases was lower than that in **Creation** cases, holding everything unchanged. Third, the chance for judges to vote for the government in New Zealand and Australia was higher than that in Canada, holding everything unchanged. Fourth, the chance for judges to vote for the government when liberal governments were in power was lower than that when conservative governments were in power, holding everything unchanged. Fifth, the chance for liberal government-appointed judges to vote for the government when liberal governments were in power was higher than all the other three possible interactions between the

variable representing the governments were in power when cases were decided and the variable representing what governments appointed the judges, holding everything unchanged. Sixth, the chance for judges to vote for the government increased as the judges accumulated experience on the bench, holding everything unchanged. Figure 1 graphs the coefficients of the three models with confidential intervals as presented in Table 5.[4]

Figure 1: Visualization of what can be learnt from the three models
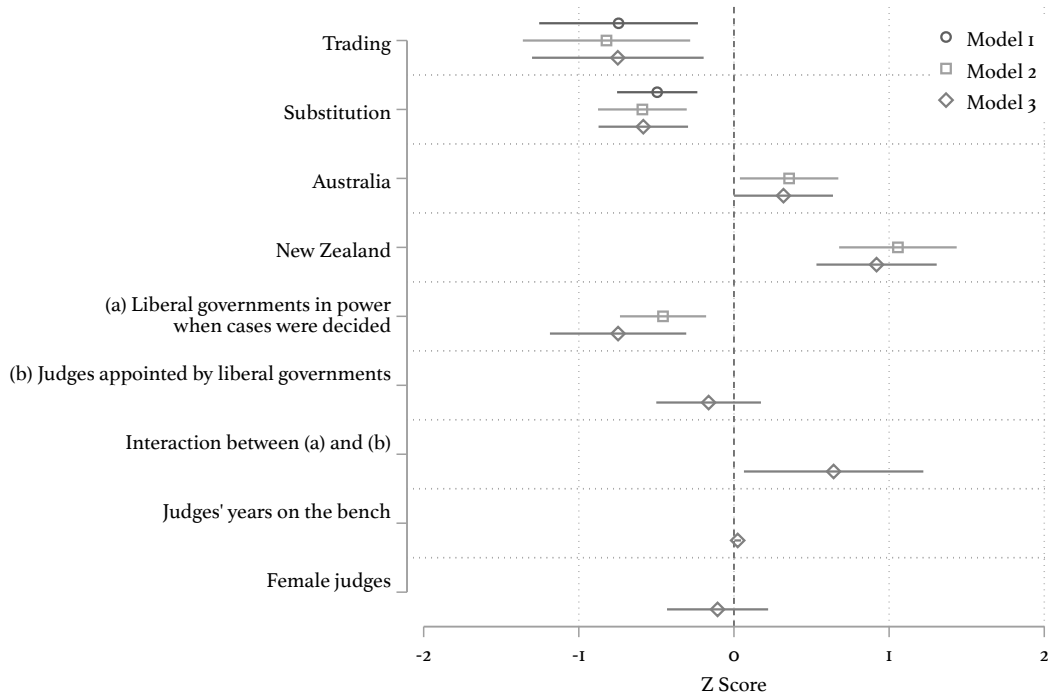
Visually speaking, the magnitude of the effects of the variables in Model 3 looks similar to that in Model 1 and Model 2, except for the variable on whether cases were decided when liberal governments were in power – the said coefficient in Model 3 is smaller than that in Model 2 at a glance. Figure 1 helps crystalizes the contributions of the following that might be positive influences of the judges' propensity to vote for the government – **Creation** cases, cases to be decided in Australia and New Zealand, cases to be decided when conservative governments are in power, cases to be decided by liberal government-appointed judges when liberal governments are in power, and cases to be decided by judges who have sat on the bench for a long time. Using the results of Model 3, a set of predicted probabilities of hypothetical judges to vote for the government was simulated under different situations. Figure 2 visualizes the simulation.

---

[4] In Figure 1, the Z Scores to the left of 0 shows that the variable at hand decreases the probabilities for judges to vote for the government in comparing with the variable used as the comparative basis. The Z Scores to the right of 0 shows that the variable at hand increases the probabilities for judges to vote for the government in comparing with the variable used as the comparative basis.

Figure 2: What-if scenarios based on Model 3

Predicted probabilities for judges to vote for the government (%)

Hypothetical judges' years on the bench
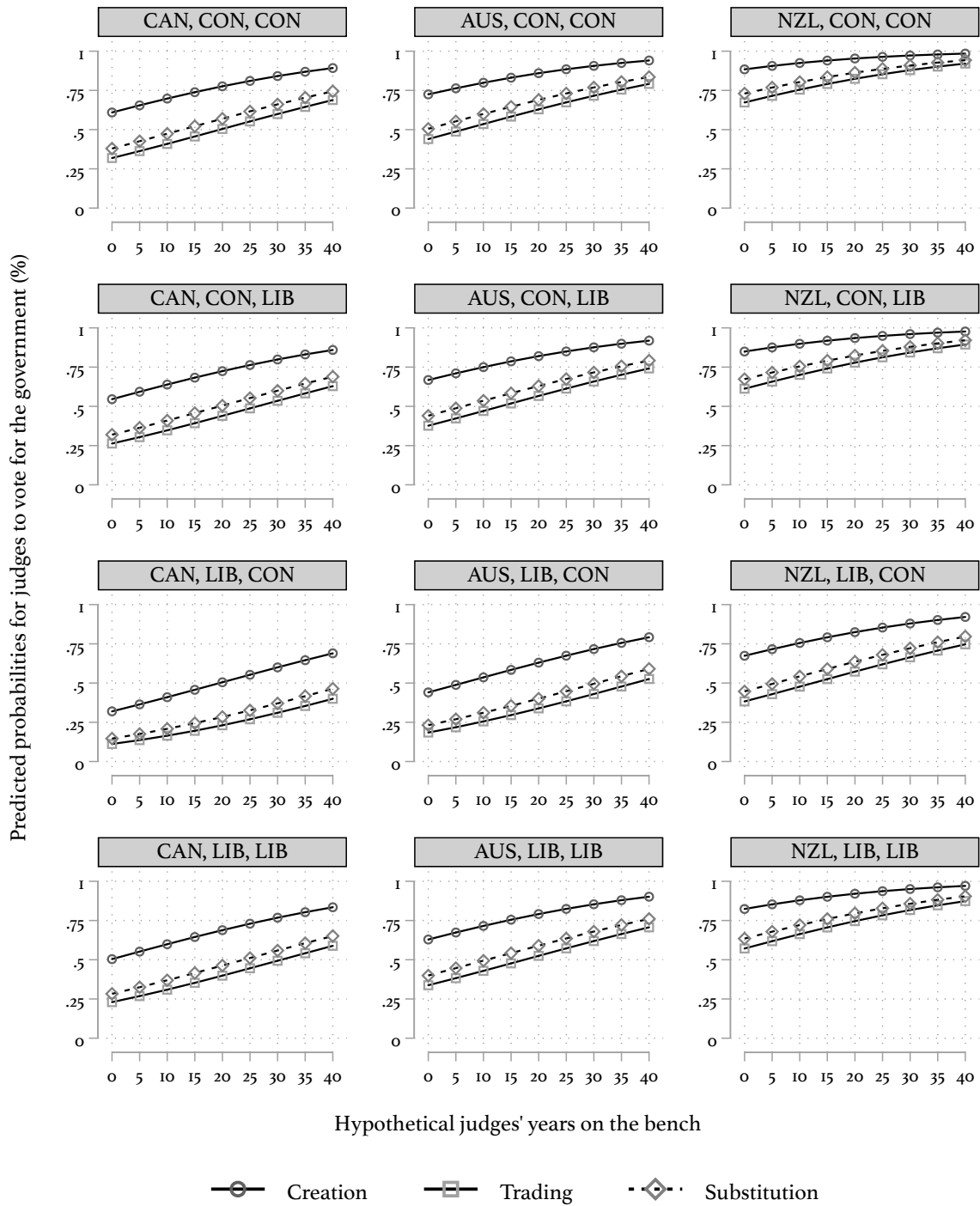
Creation   Trading   Substitution

Figure 2 displays the simulated predicted probabilities of hypothetical judges voting for the government in the long run in different situations in a 4 x 3 matrix of charts. Each chart in the matrix has a title with three labels in all caps, and the ordering of the labels from left to right in the title sums up the particular situation the chart depicts. The label on the left specifies the country (CAN as in Canada, AUS as in

Australia or NZL as in New Zealand), the label in the middle specifies the political climate (cases decided when conservative (CON) or liberal (LIB) government was in power), and the label on the right specifies the political leanings of the judges (judges appointed by conservative (CON) or liberal (LIB) governments). Each chart has three connected lines, tracing the trajectories of the predicted probabilities for the hypothetical judges to vote for the government in cases with the three transaction types over a hypothetically long career on the bench. The legend for the connected lines by case types is located at the bottom of Figure 2. The y axis of each chart indicates the probabilities, while the x axis of each indicates the number of years up to 40 years of judicial experience.

In sum, each column of charts covers four situations for one of the three countries, each row of charts covers one situation for the three countries respectively, and each of the charts shows a particular combination of the political climate when cases are decided and the political leanings of hypothetical judges who decide the cases. On the top row, the top left chart has a title of CAN, CON, CON, and it means the chart is about the simulated predicted probabilities of hypothetical Canadian judges appointed by conservative governments voting for the government when they cast their votes while conservative governments are in power at different points of their four decades on the bench. The only differences among the titles of the three charts on the top row are the countries so all three charts on the top row depict hypothetical situations of conservative government-appointed judges deciding GAAR cases when conservative governments are in power for the three countries respectively. The second row from the top depicts the hypothetical situation for Canada, Australia and New Zealand respectively when cases are decided with conservative governments in power by judges appointed by liberal governments at different points of their career. The third row from the top depicts the different situations of the three countries when cases are decided with liberal governments in power by judges appointed by conservative governments at different points of their career. The bottom row depicts the different situations of the three countries when cases are decided with liberal governments in power by judges liberal governments at different points of their career.

Figure 2 visualizes the following three simulated patterns. First, the hypothetical judges are projected to be more likely to vote for the government in **Creation** cases, followed by **Substitution** cases and then **Trading** cases at the same point in their careers. Second, the hypothetical New Zealand judges are projected to be more likely to vote for the government than the hypothetical Australian judges, who are projected to be more likely to vote for the government than the hypothetical Canadian judges at the same point in their careers. Third, the hypothetical conservative government-appointed judges are projected to be less likely to vote for the government than liberal government-appointed judges when liberal governments are in power at the same point in their careers.

Debates on the role of non-legal factors in judicial decision making never abate. Some may adhere to the tradition that cases are decided solely based on the law, while some may hold the views that cases are decided based on the law plus everything else in the world in which the cases are decided, and some may take up a position somewhere in between the two ends. Figure 2 shows the simulation of the influences of the three transaction types in judicial decision making in the three countries when governments of different ideologies are in power and when judges are appointed by governments of different ideologies, restating the questions about certainty and predictability of result in GAAR cases. Figure 2 appears to be tilted on the page. But that is an optical illusion. What the human eyes see is the upward-sloping curves tilting from lower left to upper right. Even after we are told that the charts are set straight from top to bottom on the page, we might still see what we see. One point to ponder might be that we see what we see not only just based on what it is in front of us but also based on how it is interpreted inside us.