THE US SUPREME COURT AND THE ‘AFFORDABLE CARE ACT’: AN EXERCISE IN CLOSED-DOOR FORECASTING

LEIGHTON VAUGHAN WILLIAMS

DEC 2014
DISCUSSION PAPERS IN ECONOMICS

The economic research undertaken at Nottingham Trent University covers various fields of economics. But, a large part of it was grouped into two categories, Applied Economics and Policy and Political Economy.

This paper is part of the new series, Discussion Papers in Economics.

Earlier papers in all series can be found at:

http://www.ntu.ac.uk/research/research_at_ntu/academic_schools/nbs/working_papers.html

Enquiries concerning this or any of our other Discussion Papers should be addressed to the Editors:

Dr. Marie Stack, Email: marie.stack@ntu.ac.uk
Dr. Dan Wheatley, Email: daniel.wheatley2@ntu.ac.uk

Division of Economics
Nottingham Trent University
Burton Street, Nottingham, NG1 4BU
UNITED KINGDOM.
The US Supreme Court and the ‘Affordable Care Act’: An Exercise in Closed-Door Forecasting.

Abstract
Forecasting voting outcomes has a long history, particularly in the case of US presidential elections and papal elections. The difference between these types of voting is that the former is open to a very large number of electors and to public scrutiny while the latter is closed, in the sense of a small number of decision-makers whose choices are shrouded in a layer of secrecy. This paper examines the second type of decision-making, using the 2012 US Supreme Court consideration of the Affordable Care Act to assess the relative efficacy of different methods of forecasting the outcome of these closed decisions in advance of their declaration. While the success rate of each of these methodologies has been somewhat patchy, the degree and detail of information provided by an examination of the various forecasts and the actual outcome of this case provides insights which might usefully improve forecasts of future decisions.

Introduction
Forecasting elections with a view to betting on the outcome has a long history, which is well-documented in the particular case of presidential elections in the US, and which can be traced back to 1503 for papal elections. The difference between these two types of election and decision-making is that the former is open to a very large number of electors and to public scrutiny, while the latter is closed, in the sense of a small number of decision-makers the choices of whom are shrouded in a layer of secrecy.

This paper examines the second type of decision-making, using the case of the 2012 US Supreme Court decision on the Affordable Care Act (sometimes known as ‘Obamacare’) as a case study of the relative efficacy of different methods of forecasting the outcome in

1 “All voting is a sort of gaming … and betting naturally accompanies it.” (David Thoreau, 1848, quoted in Thoreau, 1967, p.36).
advance of the declaration of the decision. This case is chosen for examination because it has features which make it particularly valuable in helping arbitrate between various hypotheses as to the best way to forecast Supreme Court decisions. In particular, the various ways used to forecast Court rulings, namely statistical methods, models and expert judgements, as well betting and prediction market, apply to this case. Moreover, there is a level of detail relating to each which is unusual, and key variables underpinning the various current forecasting models apply in this case in a way which helps discriminate the contribution of each.

More generally, the forecasting of closed-door decisions of this kind poses quite different problems to the forecasting of open elections, in which the use of scientific polls are widely used to help gauge opinion and thus help predict the eventual outcome. A key difference is that opinion polls normally ask respondents to reveal their preferences, rather than to provide a forecast. In the case of closed-door decisions, it is questionable how far the personal preferences of respondents will correlate, causatively or otherwise, with the votes of the decision-makers. If the survey is of public estimates about the outcome of the decision, rather than outcome preference, there is the problem that the public will often know very little about the attitudes or even identity of the decision-makers.

Of course, if it could be established that a link does exist between the results of surveys of public opinion and decisions made behind closed doors, this could provide a potentially fruitful avenue of research. Although some evidence does exist of a link between public opinion or the ‘policy mood’ (Stimson, 1991, 2004) and Supreme Court decisions, at least for non-salient cases (e.g. Casillas et al. 2011), the usual dependent variable here is the general ideological direction of the Court, and changes in this over time. The value of this perspective in forecasting specific cases, and in particular the highly salient Affordable Care Act (ACA) case, is much less clear or defined.

In the current absence of an established predictive link between surveys of public opinion and the outcome of particular Supreme Court decisions, therefore, the focus of this paper is the use of other options found in the forecaster’s tool kit.
In Part 1, the aim is to examine the way in which statistical methods, models, expert judgements, and prediction markets have been used to forecast the outcome of closed-door decisions prior to the 2012 health care case, and to assess their relative effectiveness.

In part 2, we shall apply the same perspective to the 2012 Supreme Court decision relating to the Affordable Care Act of 2010.

We shall use the 2012 case to compare and contrast the relative performance of each of these forecasting methodologies.

1. Forecasting US Supreme Court decisions

1a. Statistical methods and models

It might seem reasonable to suppose that the decisions that will be announced by Supreme Court justices might in some ways be predictable based on the questions they ask, and the comments they make, during oral arguments. Black et al. (2011) test this in the context of emotionally charged word choice, utilizing the Dictionary of Affect in Language (Whissell, 1989)\(^2\) to provide an objective measure of the emotional content of language. In so doing, they are drawing upon the idea that individuals most often use emotive language when particularly concerned about the outcome of the decision-making process (Zeelenberg et al, 2008). A parallel argument is that it is possible to predict the outcome of cases based on the “tenor of the argument” (Greenhouse, 2008).\(^3\) Supporting evidence is also found in Shullman (2004) and Wrightsman (2008) that justices’ language may be related to Court decisions. In particular, Black et al. examine outcomes in all orally argued cases from 1979 to 2008, using voting behaviour at individual justice level from 2004 to 2008. Allowance is also made for the potential influence of the number of questions posed to each side. In a study of 28 cases by Roberts (2005), for example, it was found that 86 per cent of the time the party receiving the most questions from the

\(^2\) See also Plutchik, 1994; Russell, 1978.

justices lost the case. A larger-scale study by Epstein et al. (2010) reached a similar conclusion. In a sample of 2,952 cases between 1979 and 2007, they found that the petitioners\textsuperscript{4} to the Court won 62 per cent of the time, the respondents 38 per cent. If the total number of words directed at the petitioner was less than at the respondent, however, the win rate of the petitioner rose to 72 per cent against 28 per cent for the respondent. If the number of words directed at the petitioner was more than at the respondent, the win rate fell from 62 per cent to 50 per cent, compared to an increase from 38 per cent to 50 per cent for the respondent. The corresponding figures for the win rate of petitioner and respondent based on number of questions (rather than number of words) showed a similar pattern, the corresponding statistics being 62 per cent (rising to 71 per cent) for the petitioner, and 38 per cent for the respondent (falling to 29 per cent) when the petitioner was asked fewer questions, and 62 per cent (falling to 53 per cent) for the petitioner when asked more questions, compared with 38 per cent (rising to 47 per cent) for the respondent.\textsuperscript{5}

The key finding of Black et al. is that when justices use unpleasant language during oral arguments, the side that suffers a great proportion of this language is more likely to lose its case. Importantly, this may also provide information on the susceptibility of the individual justices to persuasion in future cases.

Martin et al. (2004) employed a statistical forecasting model based on a handful of easily observable general case characteristics\textsuperscript{6} derived from past Supreme Court decisions, which disregarded information about the specific law or facts of the case. The model, which took into account all cases that the Court had decided prior to the 2002 term, was found to correctly predict 75.0 per cent of case outcomes during the October 2002 term.

\textsuperscript{4} The petitioner is the party that petitioned the Court to review the case. The respondent is the opponent to the petition.

\textsuperscript{5} Johnson et al. (2009) reach broadly similar conclusions in a study of 2,000 hours of oral argument and 340,000 unique questions between 1979 and 1995.

\textsuperscript{6} The case variables are circuit of origin, issue area of the case (e.g. 2\textsuperscript{nd}, 3\textsuperscript{rd}, DC, Federal Circuit), type of petitioner (e.g. US, an employer), type of respondent, ideological direction of the lower court ruling, whether the petitioner argued that a law or practice is unconstitutional.
In forecasting the votes of individual justices, the predictions were 66.7 per cent accurate.\footnote{See Ruger et al. (2004), Martin et al. (2004) for more complete description of the statistical model.}

Another method of predicting the decisions of justices is by evaluation of their ideological position. One such measure is the Martin-Quinn method\footnote{See Martin, Andrew D. and Kevin M. Quinn (2002, 2007).}, and scores derived from this method.\footnote{For a critical perspective on the value of Martin-Quinn scores, see Farnsworth, 2007; Bailey, 2012. http://www9.georgetown.edu/faculty/baileyma/CourtPref_July2012.pdf} This technique has parallels with the DW-Nominate method\footnote{See Lewis, Jeffrey B. and Keith T. Polle, 2004; Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole and Howard Rosenthal, 2009. Also: http://voteview.com/dwnomin.htm} that estimates the ideology of members of Congress based on their voting records.\footnote{See also Corey Rayburn Yung (2010).}

Martin-Quinn scores dating back to 1937 are charted in the figure below.\footnote{http://commons.wikimedia.org/wiki/File:Graph_of_Martin-Quinn_Scores_of_Supreme_Court_Justices_1937-2011.png}

Martin et al. (2004) also assessed the performance of legal experts in predicting court decisions. Their ‘experts’ consisted of 83 individuals who had “written and taught about, practiced before and or clerked at the Supreme Court.” (p. 763). Each expert was asked to predict a case within their areas of expertise. No communication was allowed between experts, and forecasts were required before oral argument. They found that the expert predictions had an accuracy rate of 59.1 per cent, compared with a 75 per cent rate for the model, but the experts did marginally better (67.9 per cent accuracy) than the model (66.7 per cent) in predicting the votes of individual justices.

The model’s relative success in this respect, argue Ruger et al. (2004), can be traced to its ability to predict more accurately the important votes of the ‘swing’ justices, Kennedy and O’Connor.

Dividing the experts into ‘legal academic’, ‘practicing attorney’, ‘former Supreme Court clerk’, ‘former Supreme Court non-clerk’ and clerk to a currently sitting Justice, they found that legal academics did worst (53%), followed by clerks to sitting Justices (54%), former Supreme Court non-clerks (57%), former Supreme Court clerks (61%), and 92% for (the very small sample of) attorneys. Because of the sample sizes and the variation in
the cases considered, it’s difficult to draw significant conclusions about the relative performance of these sub-categories of expert, but what is interesting is that the experts who had clerked or were clerking at the Supreme Court showed no evidence of outperforming the experts as a whole.\(^\text{13}\)

The statistical model also did particularly well in forecasting ‘economic activity’ cases (87.5% to 51.3%), while the experts did better than the model only in the ‘judicial power’ cases (73.7% to 50%). On issues of ‘civil rights’ and ‘criminal procedure’, the model outperformed the experts by about 20%, and marginally out-performed the experts on ‘federalism’ cases.\(^\text{14}\)

1b. Prediction markets

\(^{13}\) http://www.washingtonpost.com/blogs/ezra-klein/files/2012/06/scotuschart.jpg

\(^{14}\) http://www.washingtonpost.com/blogs/ezra-klein/files/2012/06/scotuschart2.jpg
Betting on the outcome of political outcomes has a long history,\textsuperscript{15} which is well-documented in the case of presidential elections in the US (see Rhode and Strumpf, 2004, 2008, 2014), demonstrating a notable degree of forecasting accuracy. Betting on papal elections can be traced as far back as the early sixteenth century. The difference between these two types of election and decision-making is that the former is open to public scrutiny while the latter involves closed deliberations. To this extent, Supreme Court decisions are closer in type to papal elections, about which the accuracy of the betting markets has been scrutinized elsewhere (e.g. Baumgartner, 1985, 2003; Hunt, 2012). There is limited published research, however, into the forecasting accuracy of betting markets with respect to Supreme Court decisions or appointments, and what does exist is not altogether favourable. In particular, the real-money prediction market, TradeSports, was showing, before the nomination of Judge John Roberts, an implied probability of 80 per cent that Judge Edith Clement would be appointed).\textsuperscript{16} Until “roughly two hours before the official announcement, the market was more or less completely ignorant of the existence of John Roberts, the actual nominee.”\textsuperscript{17}

Prediction markets might be particularly susceptible to failure in the case of events that are dependent on the decisions of a single person and very small number of people. Despite this, there has existed since October, 2009, a crowdsourced\textsuperscript{18} fantasy league (or ‘prediction market’), FantasySCOTUS.net, designed for “jurisprudential speculation” (Blackman, Aft and Carpenter, 2012, p.126). During the October 2009 Supreme Court Term, over 5,000 members made more than 11,000 predictions for all eighty-one cases decided. Based on these data, Blackman et al. report that the market correctly predicted the outcome in more than fifty percent of the cases decided, and that the top-ranked predictors forecast 75 percent of the cases correctly. The FantasySCOTUS market rewarded participants with bragging rights (the ‘golden gavel trophy’) rather than monetary profit.

\textsuperscript{15}“All voting is a sort of gaming … and betting naturally accompanies it.” (David Thoreau, 1848, quoted in Thoreau, 1967, p.36).
\textsuperscript{17} Id. At 120.
\textsuperscript{18} Jeff Howe, The Rise of Crowdsourcing, WIRED, June 2006, 177.
An extended perspective on the value of properly formulated prediction markets in forecasting Supreme Court decisions can be found in Cherry and Rogers (2006). The theme of their argument is that the Supreme Court is particularly suitable for a prediction market, in ways which they list. First, the number of decision-makers is limited to the nine Justices, so only nine votes need to be predicted. Further, it is possible to identify clear ideological preferences as well as past voting patterns for each of the judges. They also contend that adherence to precedent acts as a constraint on the number of possible outcomes, and that the universe of such outcomes is very limited, which makes prediction easier. They accept the potential weakness of a prediction market, however, when just one person is making a decision, such as the appointment of a Justice.¹⁹

2. Brief background to the ‘Affordable Care Act’ Supreme Court Judgement

The Affordable Care Act consists of two pieces of legislation: the Patient Protection and Affordable Care Act (PPACA), enacted on March 23, 2010, and the Health Care and Education Reconciliation Act (HCERA), enacted on March 30, 2010. Jointly they are referred to as the Affordable Care Act (ACA). Many ACA provisions went into effect immediately or shortly after the enactment of the law, such as provisions allowing young adults to stay on their parents’ plans until the age of 26 and denying insurers the right to refuse cover to children based on pre-existing conditions. Some of the more controversial aspects of the law were delayed until 2014, notably the ‘individual mandate’ on most Americans to obtain health care coverage or pay a penalty levied through the tax system. Secondly, a major expansion of Medicaid designed to cover the vast majority of Americans under 65 with incomes under 133% of the federal poverty line. In order to encourage states to comply with this policy of expansion, the Federal Government reserved the right to remove existing Medicaid funding to any states that refused to implement the expansion. Numerous challenges were made to the constitutionality of the law in courts across the country, some of which were upheld and some rejected.

¹⁹ "While information markets do an excellent job of aggregating information and making predictions, they are not mind-reading devices." (Cherry and Rogers, 2006, p.1160).
The case ended up in the Supreme Court, as Case 11-398, the Department of Health and Human Services, et al. (Petitioners) v. State of Florida, et al. (Respondents). Oral arguments began on Monday, March 26, 2012, and continued until March 28, 2012. Specifically, the Department of Health and Human Services was petitioning to reverse the judgement of the United States Court of Appeals, Eleventh Circuit, which had declared the Affordable Care Act’s individual mandate to be unconstitutional and as such not a valid exercise of Congress’ power.20

In summary, the Court heard oral argument on four separate questions, related to the Affordable Care Act: (1) is the individual mandate a penalty, or a tax, for purposes of the Anti-Injunction Act;21 (2) is the individual mandate constitutional; (3) is the mandate severable from the rest of the law; and (4) is the law’s Medicaid expansion excessively coercive upon the states?

The key issue relating to the constitutionality of the individual mandate can be separated into two parts. Is it constitutional under the Commerce Clause, which gives Congress, among other things, the power to regulate commerce among the states, and to pass all laws that are “necessary and proper” to pass these regulations? Secondly, is it constitutional under the taxing powers of Congress?

A key question in the context of the Commerce Clause is whether Congress has the right to regulate commercial activities that occur entirely within one state and never cross state lines, such as growing wheat on one’s own property. Precedent on this matter was established in the 1942 decision of Wickard v. Filburn,22 when the Supreme Court held that Congress could regulate the production of wheat on Roscoe Filburn’s farm, even if the wheat was only used for his and his family’s personal consumption, and that of his

---

20 Florida v. U.S. Dept. of Health and Human Servs., 648 F.3d 1235. 53 EBC 1649 (11th Cir. 2011) [2011 BL 210461]

21 If the mandate was a tax for purposes of the Anti-Injunction Act, the Court could not hear any challenges to it until after it was enforced, i.e. not until 2014.

22 http://en.wikipedia.org/wiki/Wickard_v._Filburn
livestock. In 2005, in the case of Gonzales v. Raich, the Court ruled that Congress could prohibit Angel Raich from using marijuana grown wholly within his state. The second key question is whether Congress has the right to regulate a person’s decision not to engage in a commercial activity, in this case a decision not to purchase health insurance. A key question in the context of the taxing powers of Congress is whether the individual mandate, which in the Act required most Americans to purchase health insurance or else pay a penalty, can be read as a tax.

2a. Using statistical methods to forecast the verdict

A study of the number of words directed, on the critical second day of oral arguments, at Solicitor General Donald Verrilli, the advocate for the Obama administration, and directed at Paul Clement and Michael Carvin, arguing against the Act, found major differences depending on the justice asking the questions.

Table 1 Number of words directed at the two sides during oral argument on the second day of the case, which covered the issue of constitutionality of the individual mandate.

<table>
<thead>
<tr>
<th>Justice</th>
<th>Verrilli (US)</th>
<th>Clement/Calvin (FL/NFIB)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kagan</td>
<td>104 (12%)</td>
<td>747 (88%)</td>
<td>851</td>
</tr>
<tr>
<td>Breyer</td>
<td>241 (12%)</td>
<td>1693 (88%)</td>
<td>1934</td>
</tr>
<tr>
<td>Sotomayor</td>
<td>310 (23%)</td>
<td>1034 (77%)</td>
<td>1344</td>
</tr>
<tr>
<td>Ginsburg</td>
<td>296 (28%)</td>
<td>752 (72%)</td>
<td>1048</td>
</tr>
<tr>
<td>Kennedy</td>
<td>434 (67%)</td>
<td>213 (33%)</td>
<td>647</td>
</tr>
<tr>
<td>Alito</td>
<td>628 (80%)</td>
<td>157 (20%)</td>
<td>785</td>
</tr>
<tr>
<td>Roberts</td>
<td>1255 (83%)</td>
<td>258 (17%)</td>
<td>1513</td>
</tr>
<tr>
<td>Scalia</td>
<td>902 (97%)</td>
<td>25 (3%)</td>
<td>927</td>
</tr>
<tr>
<td>Thomas</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>4170 (46%)</td>
<td>4879 (54%)</td>
<td>9049</td>
</tr>
</tbody>
</table>

25 State of Florida/National Federation of Independent Business
A separate calculation, undertaken here, of the number of questions/interventions by each Justice during the oral arguments reveals the following:

Table 2

<table>
<thead>
<tr>
<th>Justice</th>
<th>Verrilli (US)</th>
<th>Clement/Calvin (FL/NFIB)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kagan</td>
<td>3 (20%)</td>
<td>12 (80%)</td>
<td>15</td>
</tr>
<tr>
<td>Breyer</td>
<td>6 (21%)</td>
<td>26 (79%)</td>
<td>32</td>
</tr>
<tr>
<td>Sotomayor</td>
<td>9 (21%)</td>
<td>34 (79%)</td>
<td>43</td>
</tr>
<tr>
<td>Ginsburg</td>
<td>4 (28%)</td>
<td>14 (72%)</td>
<td>18</td>
</tr>
<tr>
<td>Kennedy</td>
<td>10 (77%)</td>
<td>3 (23%)</td>
<td>13</td>
</tr>
<tr>
<td>Alito</td>
<td>18 (90%)</td>
<td>2 (10%)</td>
<td>20</td>
</tr>
<tr>
<td>Roberts</td>
<td>24 (80%)</td>
<td>6 (20%)</td>
<td>30</td>
</tr>
<tr>
<td>Scalia</td>
<td>40 (97%)</td>
<td>3 (3%)</td>
<td>43</td>
</tr>
<tr>
<td>Thomas</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>114 (53%)</td>
<td>100 (47%)</td>
<td>214</td>
</tr>
</tbody>
</table>

It is interesting to observe here that in the aggregate 54% of the words are directed at the side opposing the constitutionality of the mandate (see Table 1). An examination of the behaviour of the likely ‘swing’ justices (Roberts, Kennedy), however, shows a different picture. 83% of Roberts’ words were directed at the Government lawyer, and 67% of those of Kennedy. This would indicate a 5-4 split in favour of striking down the mandate. Even so, the difference in word counts toward the two sides was lowest in the case of Kennedy, who is in any case the least predictable Justice (according to Epstein et al, 2009). In terms of the number of questions and interventions (see Table 2), it is interesting to note that the imbalance by Kennedy widens compared with the number of words, while Roberts’ ratio stays about the same. As for Table 1, the bottom line is an indicated 5-4 split against the constitutionality of the mandate. In both cases, of course, Thomas’ vote is inferred from his clear antipathy to the constitutionality of the mandate based on any informed reading of his prior policy and ideological preferences.

A closer examination of the transcript for Day 2^27 (see Appendix 1) reveals the use just twice of what might be termed negative, clearly emotive language, both times by Justice

[^26]: State of Florida/National Federation of Independent Business
[^27]: Day 2
Scalia in interchange with Solicitor General Verrilli. The tone and content of exchange between Chief Justice Roberts, in which Roberts asks Verrilli to clarify why Congress didn’t call the mandate a tax, is also instructive in relation to the question of whether the individual mandate can be construed as a tax for purposes of assessing its constitutionality. In that exchange, the Chief Justice allows Solicitor General Verrilli to explain why Congress used the word ‘penalty’ instead of ‘tax’, and then re-stated Verrilli’s view, without challenge.

Ryan Malphurs undertook a broader study of explicit bias in the questions posed by the individual justices, categorizing these as either “challenging, assisting or neutral.” He noted that many of the justices seemed to betray a clear bias in their questions, and sometimes an unwillingness to consider the other side. Justices can also prevent a lawyer, he argues, “from being able to articulate his or her opinion to the rest of the court to prevent other justices (from having) legitimate questions answered.”

Malphurs found that Justice Sotomayor offered more statements assisting one side — 14 for the government’s lawyer — than any other justice. An example is when she asked Paul Clement (arguing against the health law) why Congress couldn’t simply tax everyone for health care and then give an exemption for those who already have health insurance. In contrast, Justice Alito spent 91 percent of his statements challenging the administration’s case. Justice Kennedy devoted 73 percent of all remarks to challenging the law and 80 percent of all so-called challenging statements at the lawyer representing the government.

What the above analysis of the oral arguments does not do, however, is to identify points in the questioning where an antagonistic Justice (as measured by the relative number of words and/or questions directed at the two sides to the argument) provides an opportunity

---

28 See Appendix 1
29 See Appendix 2
30 For a broader insight into the use of rhetoric in the Supreme Court, see Malphurs (2013).
for the unfavoured side to clarify their case, and then to accept or reinforce the clarification without challenge. In this context, one exchange between Roberts and Verrilli is interesting (reproduced more completely in Appendix 2).

_Chief Justice Roberts to Verrilli: Why didn’t Congress call it a tax, then?_

_General Verrilli: Well - -_

_Chief Justice Roberts: You’re telling me they thought of it as a tax, they defended it on the tax power. Why didn’t they say it was a tax?_

_General Verrilli: They might have thought, Your Honor, that calling it a penalty as they did would make it more effective in accomplishing its objective. But it is - - in the Internal Revenue Code it is collected by the IRS on April 15th. I don’t think this is a situation in which you can say - -_

_Chief Justice Roberts: Well, that’s the reason. They thought it might be more effective if they called it a penalty._

This interchange between Roberts and Verrilli was clearly important in the context of the eventual decision by the Chief Justice to uphold the constitutionality of the individual mandate on the basis of Congress’ taxing power. Future research into the predictive power of oral arguments might usefully consider this additional textual instrument, i.e. identification of interchanges where a Justice measured conventionally as antagonistic to one side, offers that side an opportunity to clarify or boost their case, subsequently concluding the interchange with acceptance or reinforcement.

An alternative predictive model, based on the more sophisticated statistical approach of Bailey and Maltzman,(2011)﻿, employs the ideological preferences of the justices﻿ to

---

32 Evidence for this view is available in Auerbach et al. (2010), [http://.cbo.gov/sites/default/files/cbofiles/ftpdocs/116xx/doc11634/working_paper_2010-05-health_insurance_mandate.pdf](http://.cbo.gov/sites/default/files/cbofiles/ftpdocs/116xx/doc11634/working_paper_2010-05-health_insurance_mandate.pdf)

33 [http://prospect.org/article/will-supreme-court-overtturn-obamacare](http://prospect.org/article/will-supreme-court-overtturn-obamacare)

forecast the vote of each individual justice if ideology was the only factor guiding them. On this basis, the model predicts almost certain overturning of the law.\textsuperscript{35}

Bailey and Maltzman argue that ideology is not the only factor motivating justices, however, and that legal doctrine, notably deference to precedent, judicial restraint (deference to the legislative and executive branches) and a strict interpretation of the First Amendment’s right to free speech, are also factors in the decisions reached, though for some more than others.\textsuperscript{36}

To these the revised model adds to the 2011 model the ‘deference to precedent’ variable. Bailey and Maltzman’s 2011 model predicted that the Affordable Care Act would be struck down 5-4 based on the Court’s standard voting alignments, but when deference to precedent is added in, the forecast changes. This additional factor has a major effect on Kennedy’s likelihood to overturn (from almost certain to 46%), and a smaller effect on Roberts and Scalia. Taken together, Bailey and Maltzman predicted the likelihood of overturning the Act at just 30%.

So, including deference to precedent, Kennedy becomes the pivotal Justice. If Kennedy votes to affirm, it is possible to argue that this would increase the probability of Roberts joining him, as this would allow the Chief Justice (by virtue of siding with the majority opinion) to control the assignment of the writing of the opinion, and indeed to write the opinion himself. Bailey and Maltzman also hazarded a guess that there was some chance that Alito might follow to make the majority to affirm 7-2. To place these probabilities in more completer context, it might be useful here to include into the predictive analysis a ‘strategic’ variable, based on the opportunity which joining the majority side of the opinion gives to write to assign the opinion, including self-assigning it, as Roberts did in this case.

\textsuperscript{36} See Bailey and Maltzman (2008) for details.
Surprisingly, perhaps, they do not include the additional variable of ‘judicial restraint’. This is particularly critical in light of a statement made by Chief Justice Roberts in 2006 – “Members of Congress have been chosen by hundreds of thousands of people, millions of people. Not a single person has voted for me … And that is, to me, an important constraint. It means that I’m not there to make a judgment based on my personal policy preferences or political preferences.”\textsuperscript{37}

In light of the stated reasoning in Chief Justices Roberts’ majority opinion, read on June 28, 2012, in support of his decision to uphold the key tenets of the Act, this would seem to be a key omission in predicting the actual outcome.\textsuperscript{38}

\textsuperscript{37} Chief Justice Roberts quoted in Barnes, Robert (2006), New Justices Take to the Podium, Putting Personalities on Display, Washington Post, November 20, A15.
\textsuperscript{38} 567 U.D. (2012) 31, Opinion of ROBERTS, C.J. See Appendix 3
In either case, the problem remains for statistical models of this kind of how to assess their success. Thus election forecasts typically use the percentage of votes cast for the candidates compared to the predicted proportion. They can also rely on the standard error of the estimates in assessing a model’s accuracy. The Supreme Court, on the other hand, is made up of just nine members, which makes these assessments more challenging.

Taking the Bailey/Maltzman model as an example, and re-specifying the model, flips the predicted outcome from reverse to affirm. It also strictly speaking reverses the individual vote tally from 5-4 to reverse to 5-4 to affirm based on the revised probability of 46% that Kennedy would vote to reverse, although 46% is so close to 50% as to make the prediction of the Kennedy vote essentially a coin flip.

Another ideology-based predictive method is based on evaluation of individual justices along a liberal/conservative spectrum, using Martin-Quinn scores\(^\text{39}\) for all courts since 1937. Nate Silver (March 29, 2012)\(^\text{40}\) uses these scores to focus on the median justice along this spectrum, who he identifies as Justice Kennedy.\(^\text{41}\)\(^\text{42}\)

He notes that Kennedy scores as more conservative than any other median justice in the historic data set. Immediately to the conservative side of Kennedy is ranked Roberts, then Alito, then Scalia, with Thomas ranked as most conservative of all. On this basis, if the Health Care decision is taken purely on ideological grounds, one might expect that Roberts would only be in the majority in favour of upholding the Act if Kennedy was also of like mind. Presciently, however, Silver notes that “a few court watchers have suggested that health care could be one of the rare exceptions in which Chief Justice Roberts sides with the liberals but not Justice Kennedy.”

\(^{39}\) http://mqscores.wustl.edu/index.php

\(^{40}\) http://fivethirtyeight.blogs.nytimes.com/2012/03/29/supreme-court-may-be-most-conservative-in-modern-history/?_r=0


\(^{42}\) See also Martin, Quinn and Epstein (2005). http://mqscores.wustl.edu/media/unc05.pdf
2b. Using expert judgements to forecast the verdict

Forecasts by expert judges can be divided into those of political scientists, those of legal scholars and those of journalists in the field.

In a prediction made on November 11, 2011, political scientists Michael Bailey and Forrest Maltzman (see above) forecast that the Court would uphold the ACA by either 6-3 (Roberts, Kennedy, Ginsburg, Breyer, Sotomayor, Kagan voting to uphold) and 7-2 (with the addition of Alito). They do not explicitly break down the vote into a judgement on constitutionality under the Commerce Clause or taxing powers, but they do claim Wickard v. Filburn and Gonzales v. Raich as the “most clearly relevant” precedents. These relate to the constitutionality of Government regulation under the Commerce

Clause. Even so, they did note that Robert might possibly write “an opinion that cabins the Commerce Clause more than it is now.” Ultimately, though, Bailey and Maltzman’s prediction is based on a belief in constrained decision making. “Policy motivations,” they argued, “won’t be irrelevant, but score this one for law.”

In a prediction released on March 25, 2012, political scientist Jeffrey A. Segal forecast that the Court would strike down the ACA by 5 to 4, with Roberts, Scalia, Kennedy, Thomas and Alito in the majority. This is consistent with the attitudinal model of decision-making espoused by Segal and Spaeth (2002), whereby the decisions of Justices are driven by their personal policy preferences.

In a prediction released on March 29, 2012, political scientist Michael Evans (see above) forecast that the Court would strike down the mandate 5-4 based on the “relative number of words uttered by each Justice to the two sides regarding the constitutionality of the individual mandate under the commerce clause.”

In terms of legal scholars, a survey was conducted by ‘Purple Insights’ between May 30 and June 4, 2012, of 38 former clerks of then current Supreme Court justices, and 18 attorneys who had argued before the Court. The sample was determined to be a representative division of the populations of former clerks.

The ‘experts’ were asked to judge on a scale of zero to 100 the probability that the Court would find the individual mandate unconstitutional. Just prior to oral arguments, the average score was 57 per cent. After oral arguments, this fell to 35 per cent.

A separate survey was conducted by Bloomberg News, which in the third week of June, 2012, e-mailed questionnaires to constitutional law experts at the top 12 US law schools in US News & World Report magazine’s 2012 college rankings, asking them to predict the ruling on the constitutionality of the individual mandate. 21 responses were received.

---

46 http://themonkeycage.org/2012/03/29/predicting-the-health-care-decision-from-the-oral-arguments/
Eight thought it would be upheld, while five thought it would not, and the remaining eight judged it a ‘toss-up’.  

Legal expert opinion can also be gauged by an analysis of the American Bar Association Special Issue Preview of the case, published in 2012, which identified a “select group of academics, journalists, and lawyers, who regularly follow and/or comment on the Supreme Court. Each expert participant completed the questionnaire separately without knowing what anyone else’s predictions would be. Experts were told that their votes would be anonymous to encourage candid responses.” (p.32). The survey was apparently taken before the oral arguments stage.

These were the key findings:

**EXPERT POLL**
Is the Individual Mandate Constitutional?

**YES**
Breyer (100%) Ginsburg (100%) Kagan (100%) Sotomayor (100%) Roberts (69%)

Kennedy (53%)

**NO**
Thomas (100%) Scalia (62%) Alito (59%)

*Percentages indicate the proportion of experts polled who believe a justice will vote in a given way.

**EXPERT POLL**
Is the Individual Mandate Severable?

**YES**

---

49 In a separate Bloomberg TV round panel discussion, one of the panellists, Paula Dwyer, editorial board member of Bloomberg View, forecast that the law would be upheld 5-4, and gave the correct split of the judges. http://m.youtube.com/watch?v=d_fPw2-KZOQ
50 American Bar Association; http://www.americanbar.org/content/dam/aba/publications/supreme_court_preview/preview-healthcare.authcheckdam.pdf
Breyer (90%)  Ginsburg (90%)  Kagan (90%)  Kennedy (90%)  Sotomayor (90%)  Roberts (80%)  Alito (60%)  Scalia (60%)

NO
Thomas (90%)  

*Percentages indicate the proportion of experts polled who believe a justice will vote in a given way.

EXPERT POLL

Is the Medicaid Expansion Constitutional?
YES
Breyer (100%)  Ginsburg (100%)  Sotomayor (100%)  Kagan (91%)  Kennedy (91%)  Roberts (91%)  Scalia (78%)  Alito (73%)

NO
Thomas (73%)

*Percentages indicate the proportion of experts polled who believe a justice will vote in a given way.

EXPERT POLL

Likely Outcomes
Will the Supreme Court Uphold the ACA? YES 85%; NO 15%

What will be the Deciding Issue in the Case? Individual Mandate (91%); Anti-Injunction Act (9%).

If the Individual Mandate is found to be Unconstitutional, will the Justices Find that it is Severable from the Rest of the Act? YES 70%; NO 30%

In terms of the individual opinion of legal scholars, there are five notable examples we can consider.

On March 22, 2010, by Orin Kerr, who argued that “In the unlikely event … the Court does take [the case], I would expect a 9-0 (or possibly 8-1) vote to uphold the individual
mandate.” “… there is less than 1% chance that courts will invalidate the individual mandate as exceeding Congress’s Article 1 power.”

On August 11, 2011, Kerr issued a revised forecast that “the mandate will be upheld by a vote of anywhere from 6-3 to 8-1” (Roberts and Kennedy voting to uphold, but less sure of Scalia and Alito). The mandate would be upheld, he forecast, on the basis of deference to commerce clause precedent.

On March 23, 2010, Jonathan Adler forecast (if “forced”) that the individual mandate would survive judicial review because “Federal courts have been quite reluctant to strike down federal statutes on enumerated powers grounds for quite some time, and the individual mandate is a larger and more consequential piece of legislation than those invalidated by the Rehnquist Court. It is much easier for a court to invalidate a small piece of symbolic legislation than a major social reform.” Even so, he argued, that “while it’s a relatively safe bet to predict the Court will reaffirm federal power if pressed, the Court has confounded such expectations before – and there’ a non-trivial chance it could do so again.”

On March 28, 2010, Jack Balkin, forecast that the Supreme Court would uphold the constitutionality of the ACA under its taxing powers. “The Constitution gives Congress the power to tax and spend money for the general welfare. The tax promotes the general welfare because it makes health care more widely available and affordable. Under existing law, therefore, the tax is clearly constitutional… Opponents of the individual mandate … are really claiming that it is unconstitutional to make Americans pay taxes. The Supreme Court, however, will not be fooled, and they will reject this challenge.” Balkin, like Adler, does not provide a forecast of the split of the vote in the event that the Supreme Court took the case, but his forecast of the outcome and his analysis turned out

51 http://www.volokh.com/2010/03/22/what-are-the-chances-that-the-courts-will-strike-down-the-individual-mandate/
to be remarkably prescient in explicitly linking the Court’s likely view of the constitutionality of the mandate to the taxing power of Congress.53

On June 26, 2012, Laurence H. Tribe forecast that the Court would uphold the ACA in its entirety, based on respect for legal precedent and judicial restraint. Upholding the ACA, he argued, would offer a “partial antidote” to cases like Bush v. Gore and Citizens United and others, where the Court “has reached further than it needed … to grab onto issues that were in the middle of the political battle…”54

On June 27, 2012, David Schultz, a legal scholar and also a political scientist, forecast that the vote split would be “either 6-3 to uphold the individual mandate (Roberts majority opinion) or more likely 5-4 to strike down the individual mandate on commerce clause grounds (maybe upheld on taxing authority). I say 55-45 probability that the mandate is gone.” His reasoning was that “precedent is not in favour with this Court.”55

In terms of the individual opinions of journalists covering this area, there are four that we can consider.

On March 18, 2012, Mike Sacks forecast that the Court would uphold the ACA, under the commerce clause, based on judicial restraint and commerce clause legal precedent. “The battle this time is likely to be an intra-conservative conflict between the economic libertarianism underlying the mandate’s challenge and the traditional principles of judicial restraint that have defined right-wing jurisprudence for more than a half-century … invalidating a sitting president’s signature legislative victory on commerce clause grounds is freighted with deeply unpleasant institutional memories both for the court and the conservative legal movement.”56

---

53 http://roomfordebate.blogs.nytimes.com/2010/03/28/is-the-health-care-law-unconstitutional/?_php=true&_type=blogs&_r=1#david
54 http://video.msnbc.msn.com/jansing-and-co/47963303/#47963303
55 http://schultzstake.blogspot.co.uk/2012/06/bets-on-health-care-decision-obamacare.html
On March 21, 2012, Linda Greenhouse forecast that the Court would uphold the ACA “by a wide margin.” “If the commerce power extends to backyard marijuana growing [Gonzalez v. Raich] (as it did to backyard what growing in the famous New Deal case of Wickard v. Filburn), the notion that Congress somehow lacks the power to regulate, restructure or basically do whatever it wants in the health care sector, which accounts for 17 per cent of the gross domestic product, is far-fetched on its face.”

On March 27, 2012, Jeffrey Toobin argued that the mandate “looks like it’s going to be struck down.” He based this on his reading of the way that the oral arguments before the Court had proceeded.

On June 27, 2012, Tom Goldstein forecast that the mandate would be upheld and that the challenge to Medicaid expansion would be rejected, and predicted that Roberts would write the majority opinion.

2c. Using betting and prediction markets to forecast the verdict

The real-money prediction market, Intrade, opened a market on January 30, 2011 on whether the Affordable Care Act would be ruled unconstitutional. The market opened at an implied probability of 37%, and traded (to relatively low volumes) in that region until the oral arguments of late March, after which (to much higher volumes) it rose to in excess of 60%. By the day preceding the ruling, the market had risen to approaching 80%. Those who sold the contract at the peak won about four times the amount of their monetary exposure.

Will the Affordable Care Act be ruled unconstitutional?

57 http://opinionator.blogs.nytimes.com/2012/03/21/never-before/
59 http://www.scotusblog.com/2012/06/in-the-end/
60 The volume of trading is shown in green.
FantasyScotus.net prediction market

The forecasts offered by the Harlan Institute’s FamilyScotus.net prediction market, outlined above, were less negative about the unconstitutionality of the Affordable Care Act, and in particular the individual mandate, and were less affected by the oral arguments. The majority did predict, however, that suit would be permitted by the Anti-Injunction Act (strongly so), and also that the Medicaid expansion was constitutional and that the mandate was severable. A clear majority (though less clearly so) also believed that the individual mandate would be ruled unconstitutional. In retrospect, the issue

could have been posed as a two-part question, i.e. would the individual mandate be ruled constitutional under the Commerce Clause, and would it be ruled constitutional under the taxing powers of Congress.

### 3. The Verdict

Supreme Court verdict – 5-4 majority decision (joined by Justices Breyer, Ginsburg, Kagan and Sotomayor), written by Chief Justice Roberts:


1. The Court concluded that the Anti-Injunction Act did not bar challenges to the health care law from being ruled on immediately, instead of when the mandate would actually go into effect. The Anti-Injunction Act is a statute that says people cannot challenge a tax until after it is enforced.

“The Affordable Care Act does not require that the penalty for failing to comply with the individual mandate be treated as a tax for purposes of the Anti-Injunction Act. The Anti-Injunction Act therefore does not apply to this suit, and we may proceed to the merits.”

2. The Court ruled that the individual mandate was not permissible under Congress’s Commerce Clause power, which had been seen as the central issue of the case.

“The Framers gave Congress the power to regulate commerce, not to compel it, and for over 200 years both our decisions and Congress’s actions have reflected this understanding. There is no reason to depart from that understanding now ...The individual mandate forces individuals into commerce precisely because they elected to
refrain from commercial activity. Such a law cannot be sustained under a clause authorizing Congress to “regulate Commerce.”

Justices Breyer, Ginsburg, Kagan and Sotomayor dissented from this part of the argument.

3. The Court ruled that the Necessary and Proper Clause could not justify the mandate, either.

“Just as the individual mandate cannot be sustained as a law regulating the substantial effects of the failure to purchase health insurance, neither can it be upheld as a “necessary and proper” component of the insurance reforms. The commerce power thus does not authorize the mandate.”

Though the Court concluded that the mandate wasn’t justified under the Commerce or Necessary and Proper clauses.

Justices Breyer, Ginsburg, Kagan and Sotomayor dissented from this part of the argument.

Chief Justice Roberts added, “That is not the end of the matter.”

“The Government’s tax power argument asks us to view the statute differently than we did in considering its commerce power theory. In making its Commerce Clause argument, the Government defended the mandate as a regulation requiring individuals to purchase health insurance. The Government does not claim that the taxing power allows Congress to issue such a command. Instead, the Government asks us to read the mandate not as ordering individuals to buy insurance, but rather as imposing a tax on those who do not buy the product. The text of a statute can sometimes have more than one possible meaning. To take a familiar example, a law that reads ‘no vehicles in the park’ might, or it might not, ban bicycles in the park. And it is well established that if a statute has two possible meanings, one of which violates the Constitution, courts should adopt the meaning that does so …The most straightforward reading of the mandate is that it

---

62 Key passages of the law introducing the individual mandate, and the reference to the taxing power, is available in Appendix 3.
commands individuals to purchase insurance... But … the Commerce Clause does not give Congress that power. Under our precedent, it is therefore necessary to ask whether the Government’s alternative reading of the statute – that it only imposes a tax on those without insurance – is a reasonable one … The question is not whether that is the most natural interpretation of the mandate, but only whether it is a ‘fairly possible’ one. … As we have explained, ‘every reasonable construction must be resorted to, in order to save a statute from unconstitutionality.’ The Government asks us to interpret the mandate as imposing a tax, if it would otherwise violate the Constitution. Granting the Act the full measure of deference owed to federal statutes, it can be so read …”

In summary, the Court ruled that the penalty imposed on those who fail to purchase health insurance “looks like a tax in many respects” and that it is permitted under the previous case law of the court for a number of reasons. For example, the amount of money due is “far less than the price of insurance” and it is collected by the Internal Revenue Service (IRS) under normal means of taxation. The court noted that the mandate “is plainly designed to expand health insurance coverage” and that “taxes that seek to influence conduct are nothing new.” Moreover, the court noted, the individual mandate does not make the failure to buy health insurance unlawful. Beyond the payment to the IRS, the Court reasoned, “neither the Act nor any other law attaches negative legal consequence to not buying health insurance.”

4. The Court also ruled that Congress’s expansion of Medicaid was unconstitutional, because it threatened states with losing existing Medicaid funds if they did not agree to abide by the expansion. The Court effectively allowed the Medicaid expansion to still go forward, but only if states have the option of not participating without risking their existing Medicaid funding.

“Congress may use its spending power to create incentives for States to act in accordance with federal policies. But when ‘pressure turns into compulsion … the legislation runs

---

contrary to our system of federalism … Nothing in our opinion precludes Congress from offering funds under the Affordable Care Act to expand the availability of health care, and requiring that States accepting such funds comply with the conditions on their use. What Congress is not free to do is to penalize States that choose not to participate in that new program by taking away their existing Medicaid funding.”

*Justices Ginsburg and Sotomayor dissented from this part of the argument.*

4. Another issue was severability – what happens to the rest of the law if a part is struck down. The Court concluded that the Medicaid ruling wouldn’t affect the rest of the law.

**4. Discussion and conclusions**

Betting on the outcome of votes has a long history, and is well-documented in the case of open elections, such as US presidential elections, and closed-door elections, such as papal elections.

This paper examines the latter type of voting, using the case of the 2012 US Supreme Court decision on the constitutionality of the Affordable Care Act (‘Obamacare’) as a case study of the relative merits of different methods of forecasting the decision (statistical methods, models, expert opinion and betting or prediction markets).

There is some evidence to suggest that the relative number of words and questions directed at the different parties to the legal dispute, as well as the tone and content of the language used can serve to help predict the eventual decision. Other models use specific case characteristics or the ideological positions of the justices to help forecast the outcome.

There is other evidence to suggest that the opinions of legal experts and/or experienced Supreme Court observers can be used to help forecast the outcome of the decision-making process.

Unlike political and papal elections, there is much less evidence on the value of prediction markets (markets designed to aggregate the opinions of a diverse range of participants) or betting markets in forecasting Supreme Court and other closed-door
decisions, though in the last few years some evidence, offering a mixed message, has emerged.

In this paper, the forecasts offered by each of these methods are examined with respect to their accuracy regarding the Supreme Court decision announced on June 28, 2012 relating to the Affordable Care Act and in particular regard to the constitutionality of the Act and of the individual mandate contained in the Act.

In terms of the statistical methods for predicting the vote, it is true that more words (54% to 46%) were directed at Clement and Calvin, arguing against the Act, than at Verrilli, arguing for the Act, yet five of the justices directed significantly more words at Verrilli than at Clement/Calvin (in the case of Roberts, 83% to 17%). In terms of ideological preferences, the Bailey and Maltzman (2011) model predicted that the Act would not be upheld. Adding in a ‘deference to precedent’ variable, however, they find that the likelihood of overturning the Act would reduce to just 30 per cent, in significant part because Kennedy was now given just a 46% probability of voting to overturn the Act. Bailey and Maltzman choose not to include the other relevant variable, ‘judicial restraint’, in their model, but this omission was actually quite critical if one quite reasonably interprets Roberts’ pivotal decision on these grounds.

Even so, Bailey and Maltzman interpret the outcome as inconsistent with the claims of the attitudinal model that personal policy preferences drive Court outcomes, but instead is consistent with a combination of legal considerations and external constraints, including the Justices’ values about law and views about the appropriate role of the Supreme Court.

Even so, closer inspection of the Bailey-Maltzman analysis reveals some apparent deficiencies. In particular, Bartels argues that that they incorrectly implied that the Court would find the mandate constitutional under the commerce clause, under the ‘deference
to precedent' principle (citing in particular the precedents of Wickard v. Filburn and Gonzales v. Raich).\textsuperscript{64}

In contrast, the analysis offered by political scientists Segal and Evans was correct in their forecast of 5-4 for striking down the constitutionality of the ACA if their analysis is understood to refer to the Court’s decision on the constitutionality of the mandate under the Commerce Clause. Yet neither picked up the tax power argument that saved the individual mandate. In any case, a forecast of the individual votes which was grounded in the relative importance of ideological preferences over legal precedent and judicial restraint, would probably have generated a more accurate assessment of the actual votes of the Justices in this case, with the arguable exception of Chief Justice Roberts.

Indeed, in terms of the expert opinion of individual legal scholars, and journalists, there was some tendency to over-weight the role of legal precedent (e.g. Kerr, Tribe, Sacks, Greenhouse) and also the judicial restraint of at least some of the Justices (e.g. Adler, Tribe, Sacks), though the latter variable might have some weight in explaining Roberts’ vote.

Steven Teles develops the judicial restraint argument a little further, arguing for the attitudinal model (e.g. Segal and Spaeth, 2002) but constrained to fit a broader politics which includes deference to Congress and separation of powers. “Roberts, no doubt influenced by his position as Chief Justice, made the call that he could pull at the seam of the law pretty hard but couldn’t unravel it completely … There is an element in Supreme Court decision-making that can be explained by statesmanship rather than jurisprudence … On no really important aspect of jurisprudence did Roberts actually break from his conservative brethren, but he did make a different political judgment than they did – not on what the Court could get away with, but what was really appropriate for it to do on a matter of such great policy significance.”\textsuperscript{65}

\textsuperscript{64} http://themonkeycage.org/2012/07/19/evaluating-forecasts-of-the-supreme-courts-health-care-ruling/
\textsuperscript{65} http://www.scotusreport.com/2012/07/12/roberts-health-care-decision-statesmanship-not-jurisprudence/
This can broadly be interpreted as a concern for the institutional reputation of the Court and a desire not to have it seen as entirely politicized. Teles also develops an argument which owes more to a strategic interpretation of the behaviour of Justices, in which Justices adopt a voting strategy designed to maximize the implementation of their policy preferences in an environment where they are not the only voter. Under this interpretation, “Roberts would have joined a decision more or less striking down the mandate but severing it from the rest of the law, but he couldn’t get the rest of the four justices to go along with him. So he ended up have to cut a deal with the liberals.”

A strategic interpretation of the actions of Justices is also picked up by Brandon Bartels, who interprets the Roberts vote as consistent with the Chief Justice’s personal preference for robust federal power (as reflected in his vote on US v. Comstock), but which interprets the Commerce Clause as not permitting the individual mandate (“throwing a bone to the conservative legal movement”), while holding that the penalty for not complying with the mandate amounts to a tax, which he argues could lay the groundwork for later repeal of the ACA.

A strategic interpretation, therefore, can be seen as viewing the actions of Justices as motivated by the maximization of their personal preferences, but in the case of Roberts this might be best achieved by a decision which he wrote which not only reinforces the status and role of federal power but also limits the scope of the Commerce Clause. It also circumscribes the reach of the tax clause justification for the mandate. In particular, the penalty for non-compliance could not be punitive, and for good measure his opinion made the extension of Medicaid in the states a voluntary matter for them.

A strategic interpretation of the actions of Justices would also explain cases where Justices might switch from their ideal vote (in a one-vote Court) to join the majority

---

66 ibid.
67 http://www.scotusblog.com/case-files/cases/united-states-v-comstock/
69 http://themonkeycage.org/2012/07/19/evaluating-forecasts-of-the-supreme-courts-health-care-ruling/
opinion where this would give them the opportunity to write and shape the opinion (an opportunity in the gift of the Chief Justice, where he is a member of the majority).

Little of this seemed to be reflected in the betting on Intrade, which was more heavily driven by the perception of the flow of the oral arguments, so that the implied probability that the Act would be found unconstitutional fluctuated from generally below 40 per cent before oral arguments to generally above 60 per cent after those arguments. In the immediate run-up to the decision, the probability rose to almost 80 per cent. The FantasySCOTUS information-aggregating prediction market, in contrast, was somewhat less affected by the oral arguments, though there was a small uptick in the probability of striking down the mandate after arguments. Overall, though, the market was for several months generally implying a likelihood of between 50 per cent and 60 per cent that the individual mandate would be ruled unconstitutional.

In summary, some of the experts, notably those polled before oral arguments, performed well overall in forecasting the outcome, to the extent (in the American Bar Association survey) of identifying Chief Justice Roberts as being significantly more likely than Justice Kennedy to uphold the mandate.

The statistical methods and models offered mixed results, forecasts based on the number of questions or words directed at the parties performing poorly, in large part because of the vote of the Chief Justice. There was a special factor at play in this particular case, however, i.e. the distinction between constitutionality of the mandate under the Commerce Clause and under the taxing power of Congress. Chief Justice Roberts directed a lot of his attention in oral arguments at Solicitor General Verrilli regarding the Commerce Clause, clearly had an effect in terms of skewing the indicator on this occasion.

The Bailey and Maltzman forecast, derived from their model of constrained decision-making where ideological preferences are tempered by, for example, deference to
precedent and judicial restraint, seems (like a number of the other expert forecasts) to have overweighted the deference to precedent variable.

In terms of the stated grounds for the decision, with a couple of notable exceptions (Jack Balkin and David Schultz) there was little to suggest that forecasters were paying much attention to the possibility that the individual mandate would be upheld under the taxing powers of Congress but struck down under the Commerce Clause. With similarly few exceptions, Roberts was not seen as the most likely of the ‘conservative’ Justices to side with the ‘liberals’ (although this likelihood was clearly identified by the American Bar Association).

In terms of the prediction markets, Intrade performed poorly, especially after the oral arguments. Before the arguments, at least, the betting was indicative of the actual result. The FantasySCOTUS information-aggregating prediction market, in contrast, was somewhat less affected by the oral arguments, but for several months prior to the decision implied (though not strongly) that the individual mandate would be struck down.

In conclusion, the evidence of the 2012 Supreme Court decision about the ‘Affordable Care Act’ does not present clear evidence of the overall forecasting power of the statistical methods, models, experts or markets. By assessing the relative strengths and weaknesses of the various forecasts, however, we might be able to learn some useful lessons which can help improve our forecasts of future Supreme Court cases, in particular highly salient cases.

Most importantly, the findings produced by this study of a landmark ruling would seem to strengthen the case for applying less weight than is usual in current models of constrained decision-making to the ‘deference to precedent’ and (to a perhaps less uniform extent) the ‘judicial restraint’ variable. An explanation consistent with the evidence of this case is also that federal power is an important policy preference for Chief Justice Roberts alongside his other policy preferences. The politics of operating in a 9-member setting might also be included as a factor in forecasting the voting strategies of individual Justices who seek to maximize the realization of their policy preferences.
Wider political judgments might also play a factor, including for the Chief Justice at least some notion of the reputational integrity of the Court.

Overall, a close examination of this Supreme Court opinion would seem to strengthen the attitudinal model, albeit the preferences of Justices might perhaps be drawn more widely than is conventional in these models in defining the objectives they seek to maximize. The realization of these preferences might also be best viewed in terms of the dynamics of a nine-member Court, rather than a one-member Court, in determining how the individual Justices will act and vote in seeking to realize the most they can of their preferences. Finally, the oral argument model, which assesses likely outcomes in terms of the number of words and questions directed at each legal team should not be seen as fatally weakened by the outcome of this case, because most of the words and questions regarding the constitutionality of the individual mandate were framed in the context of the Commerce Clause, in respect of which the Court’s vote split was consistent with the predictions of the model.

More generally, it is proposed that future research might usefully seek to improve our identification of the various sources of publicly identifiable information about other types of closed-door decision, and to investigate the best way to use the range of methodologies available to the forecaster to improve the accuracy and reliability of our forecasts of these closed decisions.
References


38


Appendix 1
Emotive language:
P.35, line 20:
Scalia, responding to General Verrilli: We’re not stupid.

P.52, lines 14 to 16.
General Verrilli: It’s justifiable under its tax power.
Justice Scalia: Extraordinary.

Appendix 2
p. 48 Line 15 to p.49 Line 23:
General Verrilli: On the – December 23rd, a point of constitutional order was called to, in fact, with respect to this law. The floor sponsor, Senator Baucus, defended it as an exercise of the taxing power. In his response to the point of order, the Senate voted 60 to 39 on that proposition. The legislative history is replete with members of Congress explaining that this law is constitutional as an exercise of the taxing power. It was attacked as a tax by its opponents. So I don’t think this is a situation where you can say that Congress was avoiding any mention of the tax power. It would be one thing if Congress explicitly disavowed an exercise of the tax power. But given that it hasn’t done so, it seems to me that it’s - - not only is it fair to read this as an exercise of the tax power, but this Court has got an obligation to construe it as an exercise of the tax power, if it can be upheld on that basis.
Chief Justice Roberts: Why didn’t Congress call it a tax, then?
General Verrilli: Well - -
Chief Justice Roberts: You’re telling me they thought of it as a tax, they defended it on the tax power. Why didn’t they say it was a tax?
General Verrilli: They might have thought, Your Honor, that calling it a penalty as they did would make it more effective in accomplishing its objective.70 But it is - - in the Internal Revenue Code it is collected by the IRS on April 15th. I don’t think this is a situation in which you can say - -
Chief Justice Roberts: Well, that’s the reason. They thought it might be more effective if they called it a penalty.

70 Evidence for this view is available in Auerbach et al. (2010), http://cbo.gov/sites/default/files/cbofiles/ftpdocs/116xx/doc11634/working_paper_2010-05-health_insurance_mandate.pdf
Appendix 3

This mandate was introduced in the legislation under chapter 48, Sec. 5000A. Requirement to maintain minimum essential coverage. The key sections are part (a) the introductory paragraph and Section 1 of part (b), headed ‘Shared Responsibility Payment.’

“(a) REQUIREMENT TO MAINTAIN MINIMUM ESSENTIAL COVERAGE. – An applicant shall for each month beginning after 2013 ensure that the individual, and dependent of the individual who is an applicable individual, is covered under minimum essential coverage for such month.

(b) SHARED RESPONSIBILITY PAYMENT. –

(1) IN GENERAL. - If a taxpayer who is an applicable individual, or an applicable individual for whom the taxpayer is liable under paragraph (3), fails to meet the requirement of subsection (a) for 1 or more months, then, except as provided in subsection (e), there is hereby imposed on the taxpayer a penalty with respect to such failures in the amount determined under subsection (c).
DISCUSSION PAPERS IN ECONOMICS

2014/5  Leighton Vaughan Williams, *The US Supreme Court and the 'Affordable Care Act': An Exercise in Closed-Door Forecasting.*

2014/4  Juan Carlos Cuestas and Barry Harrison, *Unemployment Persistence in the EU15.*

2014/3  Francesco Marchionne and Sunny Parekh, *Growth, Debt, and Inequality.*


2014/1  Eghosa Igudia, Rob Ackrill, Simeon Coleman and Carlyn Dobson, *Austerity Measures or Entrepreneurial Development? The case of the Nigerian Informal Economy.*

2013/5  Barry Harrison and Theodorus Wisnu Widjaja, *Did the Financial Crisis impact on the Capital Structure of Firms?*


2013/3  Chunping Liu and Patrick Minford, *How Important is the Credit Channel? An Empirical Study of the US Banking Crisis.*


2013/1  Dan Wheatley, *Is it good to share? Debating patterns in availability and use of job share.*


2012/1  Rob Ackrill and Simeon Coleman, *Inflation Dynamics In Central And Eastern European Countries.*


2011/3  Dan Wheatley and Zhongmin Wu, *Work, Inequality, And The Dual Career Household.*

2011/2  Simeon Coleman and Kavita Sirichand, *Fractional Integration and the Volatility Of UK Interest Rates.*

2011/1  Simeon Coleman, *Investigating Business Cycle Synchronization In West Africa.*

2010/10 Vitor Leone And Bruce Philp, *Surplus-Value And Aggregate Concentration In The UK Economy, 1987-2009.*

2010/9 Robert Ackrill and Adrian Kay, *WTO Regulations and Bioenergy Sustainability Certification – Synergies and Possible Conflicts.*

2010/8 Paul Alagidede, Simeon Coleman and Juan Carlos Cuestas, *Persistence Of Inflationary Shocks: Implications For West African Monetary Union Membership.*

2010/6 Bruce Philp and Dan Wheatley, *The time scarcity and the dual career household: competing perspectives*

2010/5 Juan Carlos Cuestas, Sebastián Freille and Patricio O’Gorman, *The media and public agendas: testing for media effects in Argentina Turing the Kirchner administration*

2010/4 Vitor Leone, *From property companies to real estate investment trusts: the impact of economic and property factors in the UK commercial property returns*

2010/3 Juan Carlos Cuestas and Paulo José Regis, *Purchasing power parity in OECD countries: nonlinear unit root tests revisited*

2010/2 Juan Carlos Cuestas and Bruce Philp, *Exploitation and the class struggle*

2010/1 Barry Harrison and Winston Moore, *Nonlinearities in Stock Returns for Some Recent Entrants to the EU*

2009/7 Joao R. Faria, Le Wang and Zhongmin Wu, *Debts on debts*

2009/6 Juan Carlos Cuestas and Luis A. Gil-Alana, *Unemployment hysteresis, structural changes, non-linearities and fractional integration in Central and Eastern Europe*

2009/5 Juan Carlos Cuestas and Javier Ordóñez, *Unemployment and common smooth transition trends in Central and Eastern European Countries*

2009/4 Stephen Dobson and Carlyn Ramlogan, *Is there a trade-off between income inequality and corruption? Evidence from Latin America*

2009/3 Juan Carlos Cuestas and Luís Alberiko Gil-Alana, *Further evidence on the PPP analysis of the Australian dollar: non-linearities, structural changes and fractional integration*

2009/2 Estefanía Mourelle and Juan Carlos Cuestas, *Inflation persistence and asymmetries: Evidence for African countries*
2009/1 Juan Carlos Cuestas and Barry Harrison, *Further evidence on the real interest rate parity hypothesis in Central and Eastern European Countries: unit roots and nonlinearities*

2008/16 Simeon Coleman, *Inflation persistence in the Franc Zone: evidence from disaggregated prices*

2008/15 Juan Carlos Cuestas and Paulo Regis, *Nonlinearities and the order of integration of order prices*

2008/14 Peter Dawson and Stephen Dobson, *The influence of social pressure and nationality on individual decisions: evidence from the behaviour of referees*

2008/13 Juan Carlos Cuestas and Barry Harrison, *Testing for stationarity of inflation in Central and Eastern European Countries*

2008/12 Juan Carlos Cuestas and Dean Garratt, *Is real GDP per capita a stationary process? Smooth transitions, nonlinear trends and unit root testing*

2008/11 Antonio Rodriguez Andres and Carlyn Ramlogan-Dobson, *Corruption, privatisation and the distribution of income in Latin America*

2008/10 Stephen Dobson and Carlyn Ramlogan, *Is there an openness Kuznets curve? Evidence from Latin America*

2008/9 Stephen Dobson, John Goddard and Frank Stähler, *Effort levels in contests: an empirical application of the Tullock model*

2008/8 Juan Carlos Cuestas and Estefania Mourelle, *Nonlinearities in real exchange rate determination: Do African exchange rates follow a random walk?*

2008/7 Stephen Dobson and John Goddard, *Strategic behaviour and risk taking in football*

2008/6 Joao Ricardo Faria, Juan Carlos Cuestas and Estefania Mourelle, *Entrepreneurship and unemployment: A nonlinear bidirectional causality?*

2008/5 Dan Wheatley, Irene Hardill and Bruce Philp, "Managing" reductions in working hours: A study of work-time and leisure preferences in the UK industry

2008/4 Adrian Kay and Robert Ackrill, *Institutional change in the international governance of agriculture: a revised account*

2008/3 Juan Carlos Cuestas and Paulo José Regis, *Testing for PPP in Australia: Evidence from unit root test against nonlinear trend stationarity alternatives*

2008/2 João Ricardo Faria, Juan Carlos Cuestas and Luis Gil-Alana, *Unemployment*
and entrepreneurship: A Cyclical Relation

2008/1 Zhongmin Wu, Mark Baimbridge and Yu Zhu, Multiple Job Holding in the United Kingdom: Evidence from the British Household Panel Survey

DISCUSSION PAPERS IN POLITICAL ECONOMY

2006/3 Ioana Negru, On Homogeneity and Pluralism within Economics Schools of Thought
2006/2 David Harvie and Bruce Philp, Learning and Assessment in a Reading Group Format or Reading Capital... For Marks
2006/1 David Harvie, Bruce Philp and Gary Slater, Regional Well-Being and ‘Social Productivity’ in Great Britain’
2004/2 Massimo De Angelis and David Harvie, Globalisation? No Question: Foreign Direct Investment and Labour Commanded
2004/1 David Harvie, Value-Production and Struggle in the Classroom, or, Educators Within, Against and Beyond Capital

DISCUSSION PAPERS IN APPLIED ECONOMICS AND POLICY

2007/2 Juan Carlos Cuestas, Purchasing Power Parity in Central and Eastern European Countries: An Analysis of Unit Roots and Non-linearities
2007/1 Juan Carlos Cuestas and Javier Ordóñez, Testing for Price Convergence among Mercosur Countries
2006/2 Rahmi Cetin and Robert Ackrill, Foreign Investment and the Export of Foreign and Local Firms: An Analysis of Turkish Manufacturing
2006/1 Robert Ackrill and Adrian Kay, The EU Financial Perspective 2007-2013 and the Forces that Shaped the Final Agreement
2004/5 Michael A. Smith, David Paton and Leighton Vaughan-Williams, Costs, Biases and Betting markets: New evidence
2004/4 Chris Forde and Gary Slater, Agency Working in Britain: Character, Consequences and Regulation
2004/3 Barry Harrison and David Paton, Do ‘Fat Tails’ Matter in GARCH Estimation? Stock market efficiency in Romania and the Czech Republic
2004/2 Dean Garratt and Rebecca Taylor, Issue-based Teaching in Economics
2004/1 Michael McCann, *Motives for Acquisitions in the UK*

2003/6 Chris Forde and Gary Slater, *The Nature and Experience of Agency Working in Britain*

2003/5 Eugen Mihaita, *Generating Hypothetical Rates of Return for the Romanian Fully Funded Pension Funds*

2003/4 Eugen Mihaita, *The Romanian Pension Reform*

2003/3 Joshy Easaw and Dean Garratt, *Impact of the UK General Election on Total Government Expenditure Cycles: Theory and Evidence*

2003/2 Dean Garratt, *Rates of Return to Owner-Occupation in the UK Housing Market*