Chatham-Kent Woodlot Consultation Security Report

Background

We were asked by the Senior Administration of Chatham Kent to do an enhanced security analysis of the recently concluded Woodlot consultation process. The purpose was to assess if the results had been influenced by (a) special interest groups or (b) people residing outside of the municipality or attempting fraudulent activity.

Summary of Findings

Our report into the Woodlot consultation process indicated that there was community support for a mix of measures by the municipality including education, incentives and regulation. In terms of community sentiment, we found that there would be a small minority opposed to such steps but that the community was generally in support.

This report examines the data collected in light of the two questions above. Our conclusion, after extensive analysis, is that there were significant attempts to influence the outcome of the process, both legitimate and illegitimate. We found that the activities of suspicious users and special interest groups were primarily aimed at promoting landowner rights. However, correcting for this bias does not substantially change the final results of the survey.

Division Around Regulation

The impetus of this additional research was the report, by administration, of allegations by members of the public that special interest groups and groups operating from outside Chatham-Kent were operating systematically to undermine the process by adding extra votes to sway the results.

We understood from administration, and from our analysis of the comments and vote information, that the key dividing issue was the implementation (or not) of having a regulatory aspect to the bylaw —that is, a provision that would restrict the rights of property owners to cut down trees.

Although other steps were on the table (education, incentives), we understood that regulation, and its associated impact on woodlot owners' property rights was the key dividing issue.
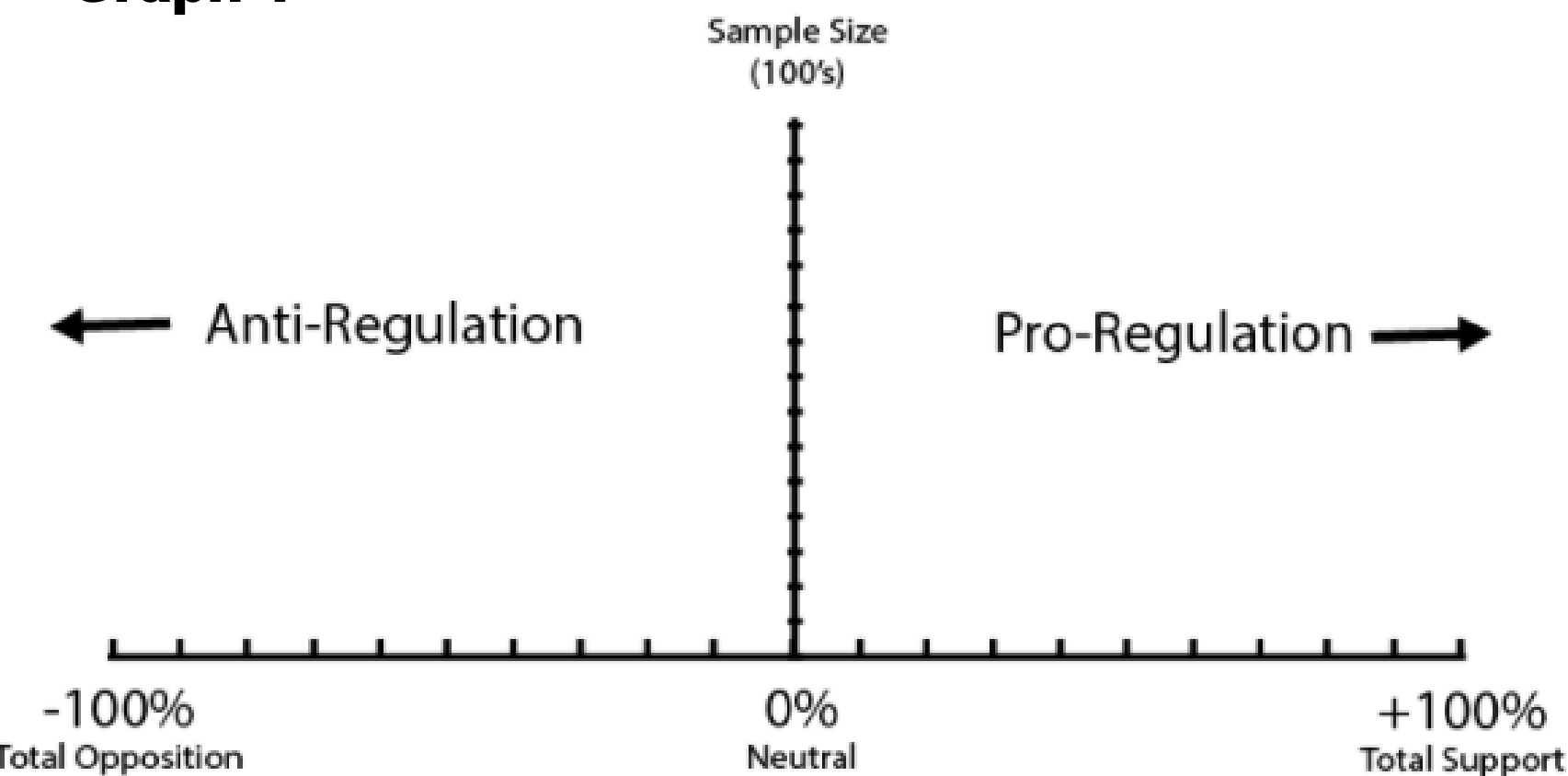
Our Approach

Graph 1

We segmented all the participants based on their support for, or opposition to, regulatory intervention to restrict woodlot owner rights. We assigned participants a "**regulation preference**" or "**RP Score**" of +1 if they were in support of regulation; and an RP Score of -1 if they were opposed to regulation.

We were able to unambiguously classify 82% of participants in this way. 13% of participants did not provide enough information to assess their preference, and 5% were ambiguous in their responses.
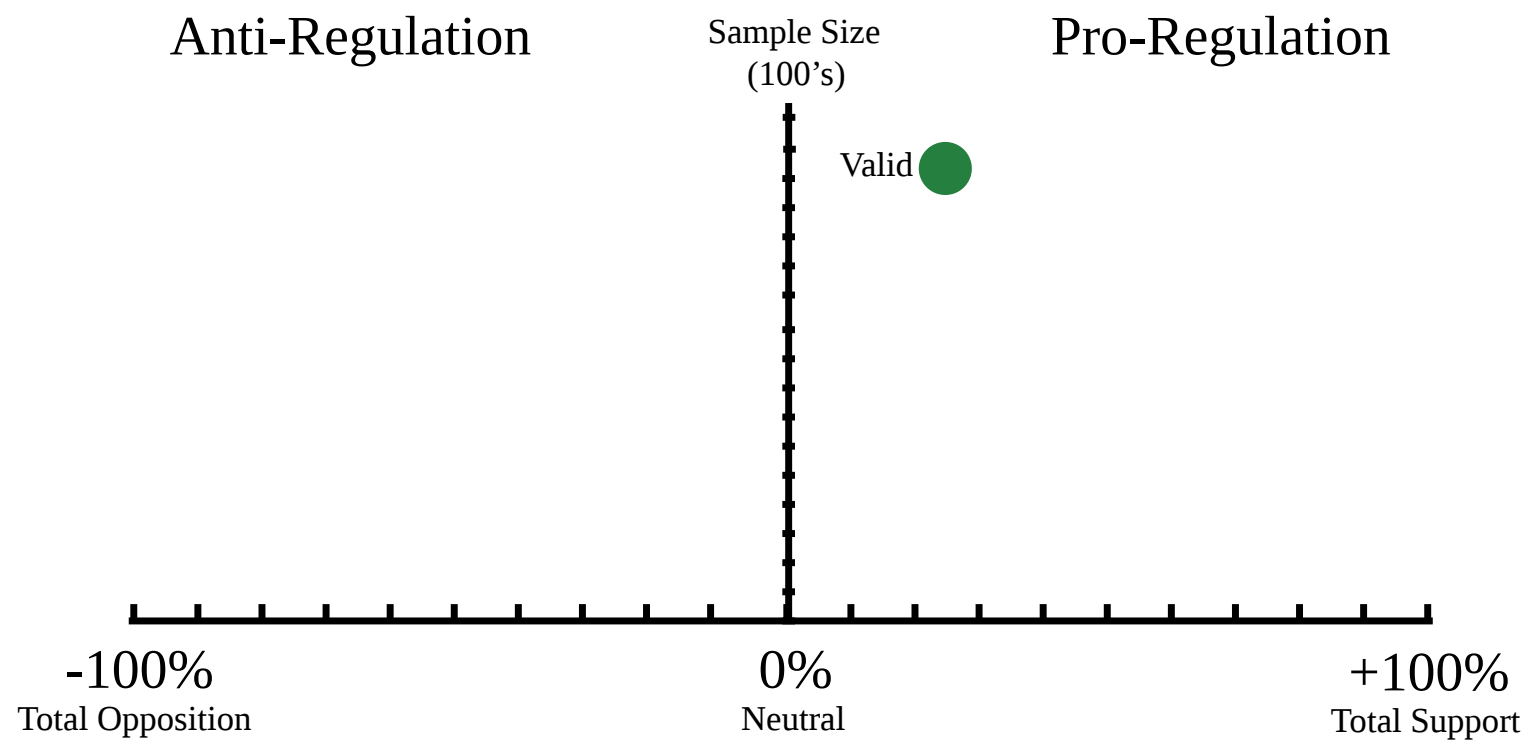
From this, we are able to arrive at an RP Score of subgroups of participants, by taking the average RP Score of the participants in that group. In other words, a subgroup with a Regulation Preference Score greater than 0 would be in favour of regulation; a subgroup with a Regulation Score less than 0 would be opposed to regulation.

Using the RP Score, we are able to locate these subgroups on Graph 1, with height of the subgroup corresponding to the size (and statistical accuracy) of the RP Score

Valid Data

Graph 2

Anti-Regulation | Sample Size (100's) | Pro-Regulation

Valid ●

-100%
Total Opposition

0%
Neutral

+100%
Total Support

We undertook a validation of all user-generated data before providing clients with their final report. Our security analysis included, but was not limited to, device IDs, IP addresses as well as attempts to obscure that information and patterns of activity on the platform. We used various techniques to determine the location of participants, as well as attempts by individuals to submit more than one entry. We also conducted analysis on user-generated data to identify patterns associated with fraudulent activity. Entries that were flagged as suspicious were excluded from the analysis and final report.
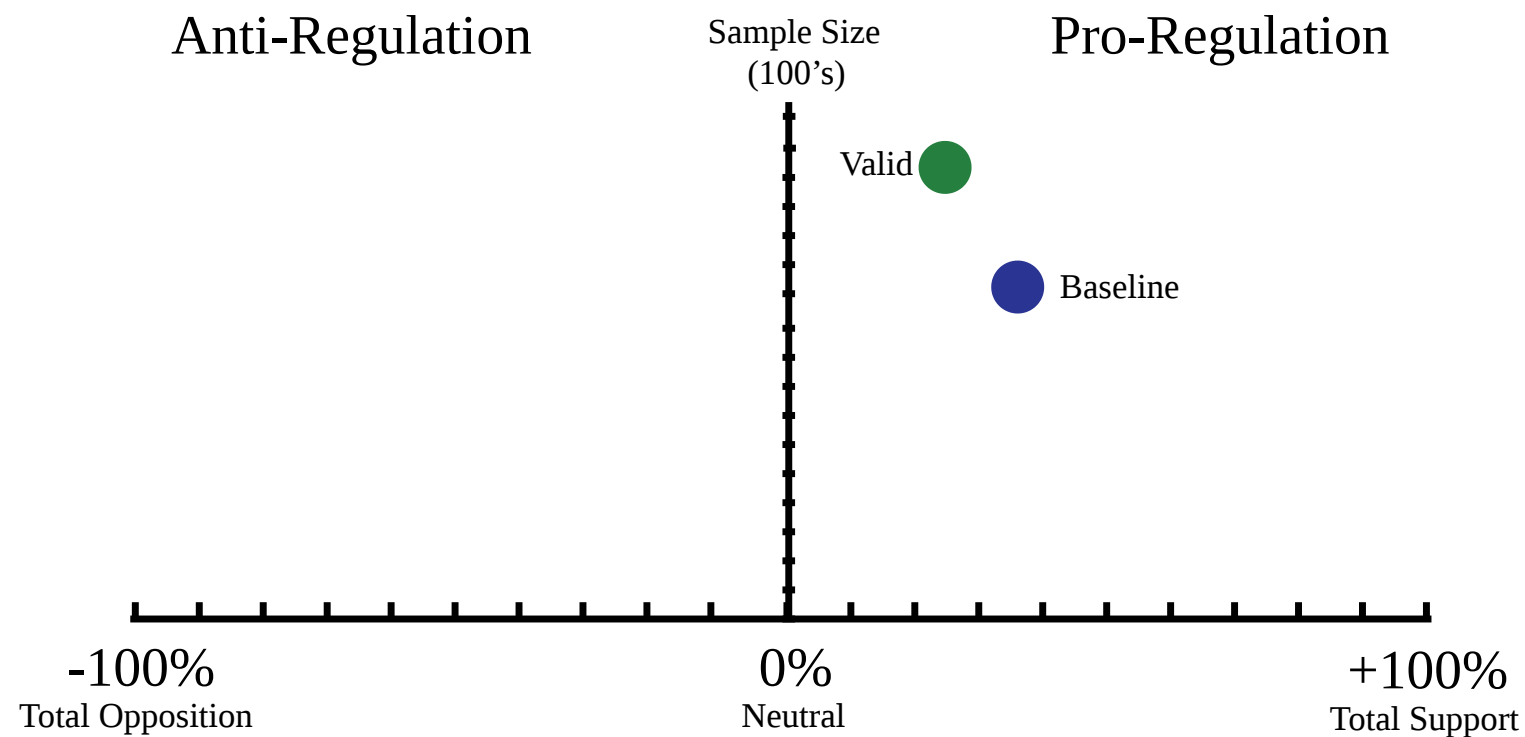
Our analysis, which included software tools as well as human analysis, flagged about 250 user accounts out of 1,763 as being suspicious.

The group of participants whose submission was deemed reliable enough to be included in the report (the "valid data") had a RP Score +26.8% . That is, they were slightly in support of regulation.

Baseline Data

Graph 3

There is always a tradeoff between security and inclusivity. Some people may be valid residents of Chatham Kent, but be participating while on a trip, or perhaps they have property in Chatham Kent but reside elsewhere. Determining who is a valid participant ultimately turns on abstract legal definitions—for example, can an environmentalist from a neighbouring community vote? Likewise, can the member of a woodlot-owning family who recently moved away vote?

We can avoid these legal questions by applying a filter that traces participation directly to a physical location within Chatham Kent. This eliminated about 400 accounts (23% of the total). While it is not necessarily "the truth", this smaller but likely higher quality dataset gives a valuable baseline for comparison.

When we applied a filter that only counted participants that could be traced directly to Chatham Kent, the RP Score increased to 36%; moderately in favour of regulation. This 10% deviation suggests that those participating from outside Chatham Kent are more likely to be against regulation.

Unreliable Data: Distance and Anonymity

There are many reasons why someone might be participating in this process from an area outside of Chatham Kent. They may be on a trip or own land in Chatham Kent while living elsewhere. This was not, in itself, enough to disqualify their submission from entry into the report. However, it is also a potential problem, if as alleged there is a group organizing support from outside the municipality of people who are not residents. Therefore, such data was not regarded as fully reliable.

Many people use VPNs as a personal security device, which was not in itself enough to disqualify participants. However, this can also signal an attempt by people to disguise their IP and hide their identity. Such data is not regarded as fully reliable. Sometimes this data appears as a large distant city.

Therefore it is useful to look at the unreliable data to determine how it compares to the Valid and Baseline data. If there is a significant deviation, either for or against regulation, then we can find evidence of a concerted campaign.

Those living more than 500km away ("Distant Users"), including those using anonymized VPNs, had an RP Score of -1%. Those living less than 500km away("Nearby Users") had an RP Score of +32%.
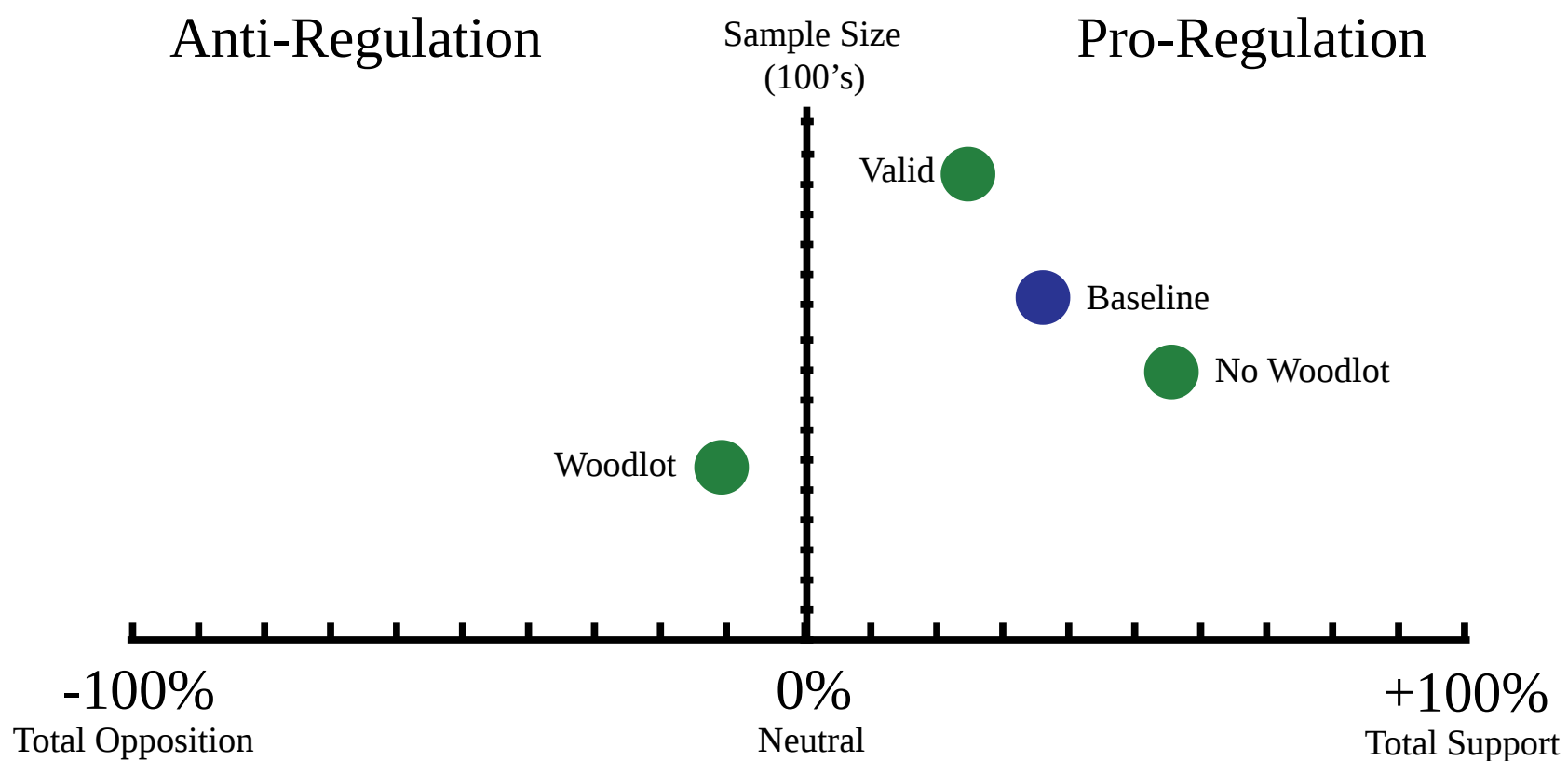
Unreliable Data: Distance and Anonymity
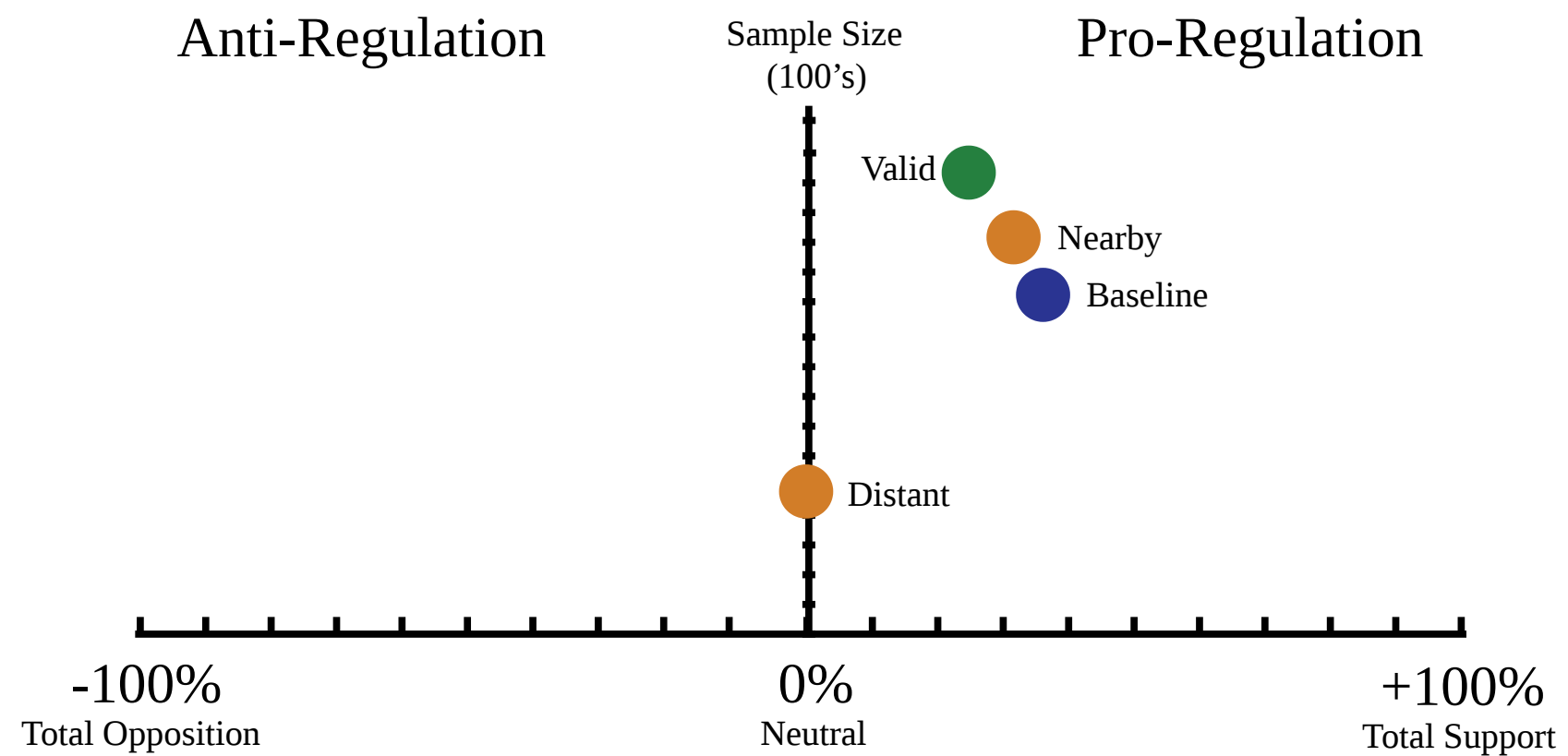
Analysis: Distant accounts were about 20% more opposed to regulation than Valid accounts. Nearby accounts supported regulation at about the same level as the Valid accounts.

Graph 4

Anti-Regulation   Sample Size (100's)   Pro-Regulation

Valid

Nearby

Baseline

Distant

-100%
Total Opposition

0%
Neutral

+100%
Total Support

Woodlot Owners and Non-Woodlot Owners

Graph 5

Anti-Regulation   Sample Size (100's)   Pro-Regulation

Valid

Baseline

No Woodlot

Woodlot

-100%   0%   +100%
Total Opposition   Neutral   Total Support

We analyzed the differences in RP Score between Woodlot and Non-Woodlot accounts, based upon answers to the question "Do you live on a woodlot property?"

When we segmented these two groups along the Regulation Preference question, we obtained the following results:

- RP Score of Woodlot accounts: -10.5%
- RP Score of Non-Woodlot accounts: 52.7%

These results point to a clear ideological divide between those that own woodlots, and those that don't; a gap of more than 60%. However, the groups are not equal in the unity of their opinion; a significant proportion of woodlot accounts were in favour of regulation, and so as a group their opposition to regulation was divided. A much lower proportion of non-woodlot accounts opposed regulation, and so the sentiment of that group was relatively consistent.
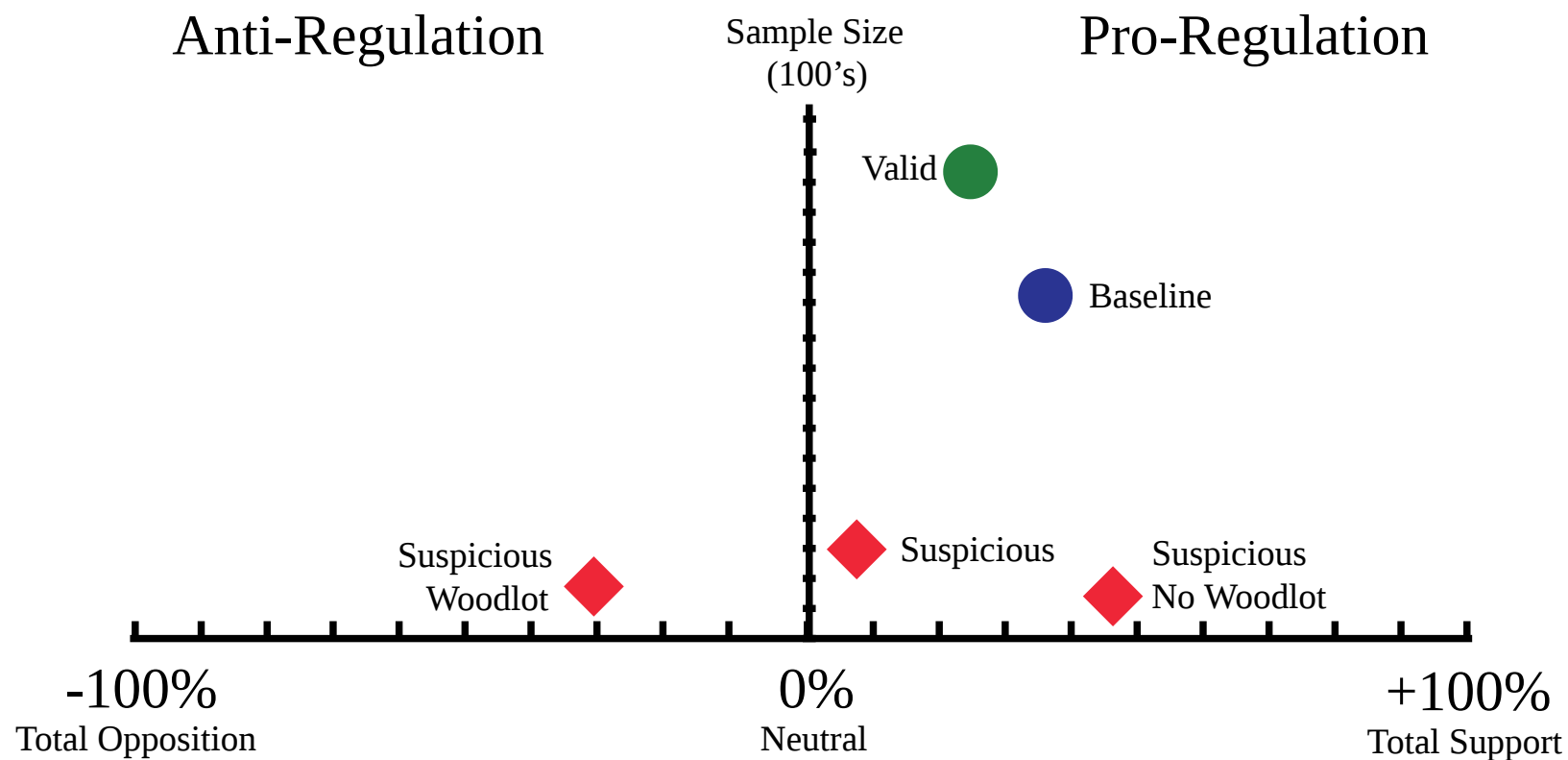
Enhanced Analysis

We undertook a deeper level of analysis to look for patterns of organized activity within the data that was identified as suspicious. We also looked for unusual levels of participation from certain segments which could correspond to special interest groups.

Suspicious Data

Graph 6

Anti-Regulation   Sample Size
(100's)   Pro-Regulation

Valid ●

Baseline ●

Suspicious ◆

Suspicious
Woodlot ◆   Suspicious ◆   Suspicious
No Woodlot ◆

-100%   0%   +100%
Total Opposition   Neutral   Total Support

We looked at the subgroup of user accounts which has been flagged as suspicious ("suspicious accounts") and not included in the report.

The RP Score of suspicious accounts was: +6.7%, that is, they were close to neutral.

Approximately half of the suspicious accounts answered "yes" to the question "do you live on a woodlot property" and half answered "no."

The RP Score of suspicious accounts that answered No to the woodlot property question was +45.5%

The RP Score of suspicious accounts that answered Yes to the woodlot property question was = -31.2%

Suspicious accounts were 20% more likely to oppose regulation than the Valid or Baseline accounts. Suspicious woodlot accounts were 50% more opposed to regulation than the Valid accounts. Suspicious non-woodlot accounts were 20% more supportive of regulation than the Valid accounts.

Conclusions Regarding Fraud

There were a significant number of attempts to undermine the data in this process. These attempts were identified and blocked during the data validation stage, and the suspicious votes provided were not included in the reports analysis.

Generally, those attempting to undermine the data were opposed to regulation; that is, they submitted votes against regulation. This was not homogenous; there were some attempts to promote a pro-regulation position. But the overwhelming majority of those attempts tried to skew the results against regulation.

If there was an organized attempt to undermine this process, that effort was undertaken with the objective of presenting the community to be more opposed to regulation than it really was.
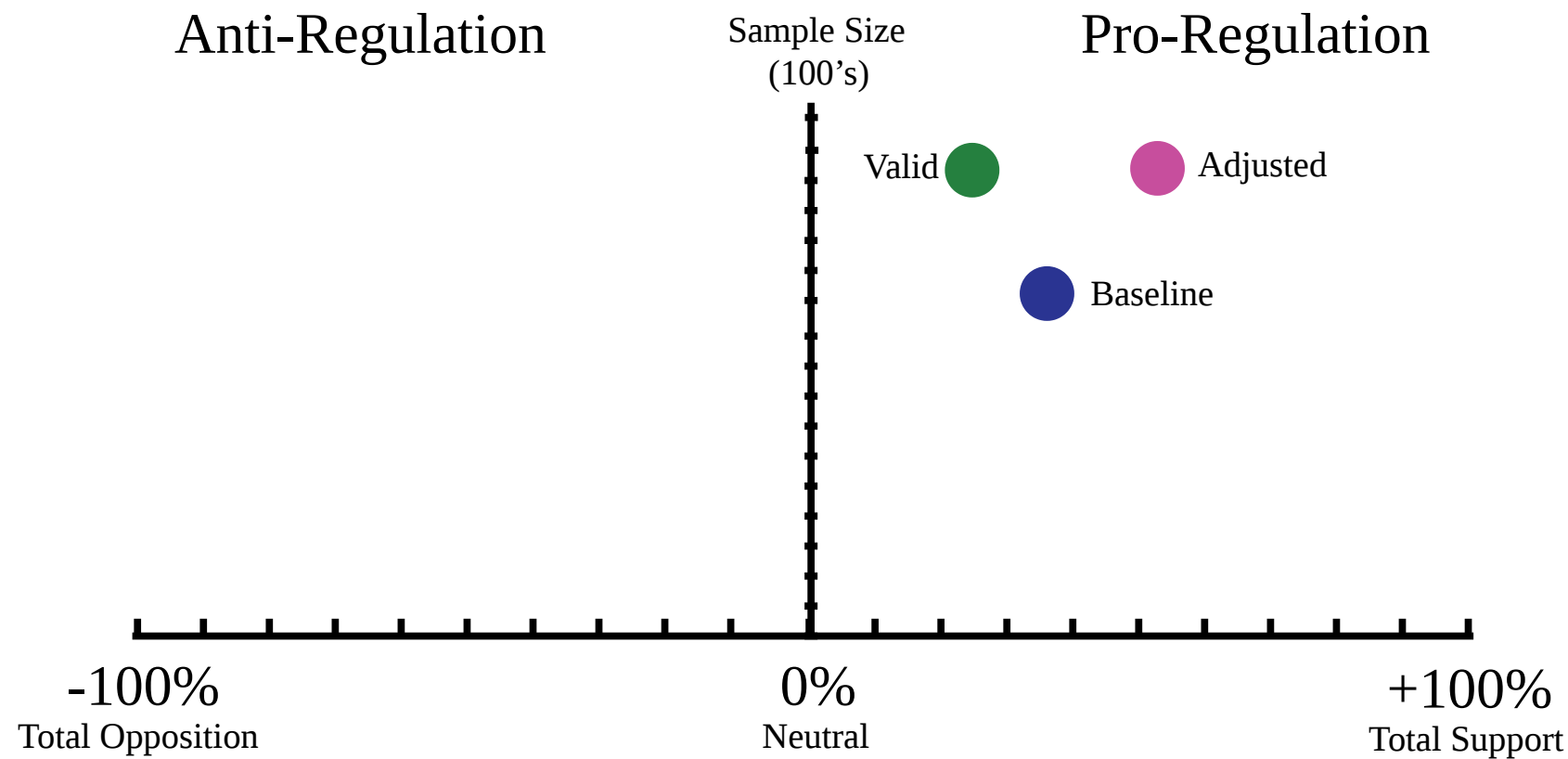
There may also have been deliberate attempts to undermine the process generally by entering "fake votes" on both sides of the question. One device in particular was used 24 times, from a variety of locations around Ontario (including VPNs), to cast votes both for and against regulation. Mostly those votes were against regulation.

Demographic Analysis

Graph 7

Anti-Regulation   Sample Size (100's)   Pro-Regulation

Valid ●   ● Adjusted

● Baseline

-100%   0%   +100%
Total Opposition   Neutral   Total Support

We were also asked to evaluate the data to determine if there was evidence of efforts by special interest groups to influence the process. For this analysis, we expanded our view to include the Valid data.

We found that 62% of participants identified as Non-woodlot users, whereas 38% of participants were Woodlot users.

We know that private woodlots are distributed across approximately 3000 separate properties in Chatham-Kent, in an area with a population of 105,529. If we assume that there are 2 adults per household, on average, then about 6,000 residents own one or more properties that feature woodlots on them which represents about 6% of the Chatham Kent population.

Therefore, woodlot accounts were overrepresented by about 600%. That is, we saw 6X more woodlot residents participating than we would if the results were generated by a random sample of Chatham Kent residents.

If we were to "correct" the sample by reweighting the votes of woodlot owners down to 6% of the total sample, then the RP Score of representative cohort would be +50%

Special Interest Groups

We found evidence that a special interest group had a disproportionate impact on the results, in this case, those participants advocating against regulation on woodlot ownership. This is not to say the impact was unfair—organizing voter turnout and ensuring high participation is a long-held tradition in democracy. However, the results are not "purely" representative of the community and were significantly affected by woodlot owners due to the self-selected nature of the sample.

If this process had been conducted by random jury or through participation of 100% of Chatham Kent residents (an unlikely scenario) then the outcome would show significantly stronger support for regulation.

Conclusions

Final Conclusions

The report provided is based on a "valid" sample of participants with an RP Score of +26%, which is slightly in favour of regulation. However, this result likely underestimates the true level of support in the community for woodlot regulation, due to efforts - both legitimate and illegitimate - to sway the results against regulation.