

IMPROVING POLITICAL REPRESENTATION THROUGH DATA

— PETER LOEWEN

A central assumption of democratic government is that representatives know important things about the citizens they represent. Successful governments wish to know what their citizens want them to do now and in the future. Politicians, political staffers, and political advisors seek to know the needs and political preferences of their voters and constituents. In some cases, politicians will seek to know the behavior of citizens, and how it is related to government policy. Central to this assumption related democratic representation is another assumption: politicians have available the information required to learn about citizens.

In this chapter, I do three things. First, I provide context for some of the on-the-ground facts which outline the gap between what we want our politicians to know and what they can know. I then briefly review the kinds of data that could improve representation. I next discuss two broad areas of political representation and public policy making that could be materially improved by a greater engagement of data by politicians and public servants. I conclude by considering how various democratic values and practices not only make it possible for public servants and politicians to engage in more systematic learning, but also confer an advantage to democracies.

It is important to acknowledge upfront that taking learning from data seriously may not be an existential crisis for democracies, but, regardless, it is a major one. At no time since the end of history¹, i.e. the end of the cold war, has there been such a contest of systems as there is today. Conflict in Ukraine underscores a fragile

European political system. Populism and antisystem sentiment represent a breaking down of the trust required to make democratic delegation work. In contrast to these faltering systems, and no matter how bumpy recent times have been, China's ascent is even more remarkable. Underwriting much of China's governance success is an unyielding commitment to knowing what its citizens are doing, what they are thinking, what they care about, and how well they are being served by their local and regional governments.² Systems of constant surveillance and social crediting fundamentally and negatively change the relationship between citizens and their state. But if one side of the coin is a total surveillance state, the other side is a belief that governments should know as much about what citizens want and think as possible so that that government can do its job better. Arguably, the problem with China is not the ambition of knowing everything about its citizens, but rather wishing to use this information to assert control over citizens rather than democratically respond to free citizens. In this, China is not alone among authoritarian states, just ahead of its contemporaries in how much progress the state has made and control China has been able to assert over its citizens.

What if democratic states committed themselves to knowing just as much in the aggregate about their citizens, while using this information democratically, in a manner that was respectful both of privacy and the democratic liberties of citizens?

WHAT DO POLITICIANS KNOW ABOUT CITIZENS? WHAT DO THEY KNOW ABOUT POLICY?

A classic, if simplified, way of thinking about the roles of politicians in a democracy is to classify them as delegates or trustees.³ Delegates believe that their central function is to deliver in government the policies that citizens want. Trustees, by contrast, are not so interested in doing what citizens want as they are in delivering on what citizens need, even if citizens cannot easily articulate (or know) those needs. In this framework, good delegates will know what citizens want them to do (i.e. their preferences) and good trustees will not know about the facts of their voters' lives or what their constituents need from government.

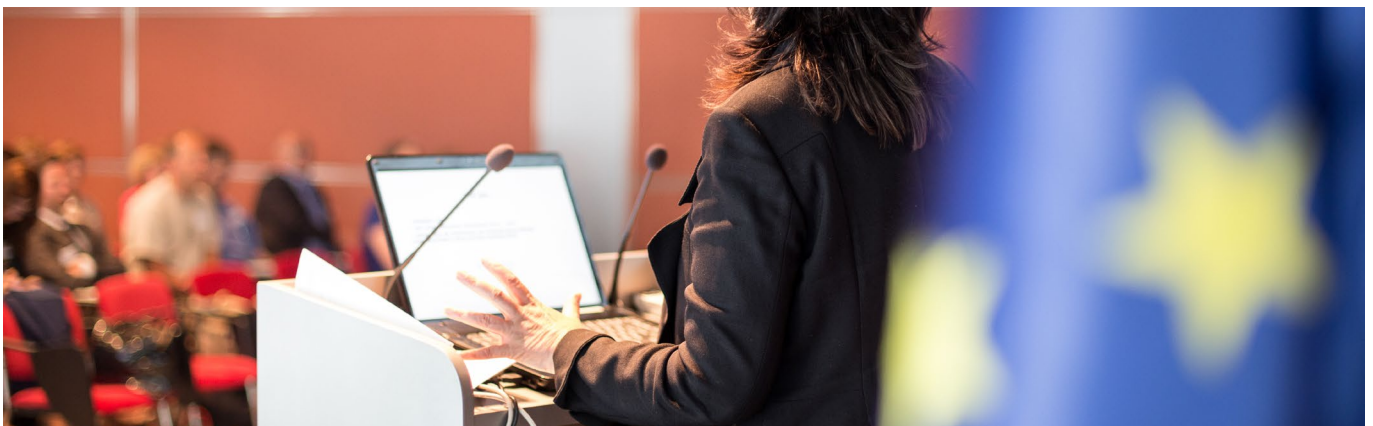
On the knowledge side, a growing body of evidence suggests that politicians in democratic countries are often systematically biased or incorrect in their perceptions of what citizens want. This has been well-documented in the United States, where a clear conservative bias exists in politicians' perceptions of citizens' preferences.^{4,5} But it exists elsewhere, too. For example,⁶ show that politicians have similar conservative biases in five other countries, and ⁷ show that politicians in several countries are not even all that good at identifying on what side of an issue a majority of citizens fall. On balance, the evidence suggests that politicians do not know what citizens want across a variety of issues. Other studies suggest that they might not be all that interested in even learning about citizen preferences ⁸, perhaps because it is so difficult to access timely and relevant data. This is contrary to much earlier evidence which suggests strong linkages between constituency preferences and politicians'

actions.⁹ Whether politicians are in fact getting worse at knowing citizens' preferences is beside the point, largely. The bottom line is that they do not know what citizens want as well as we might expect or as much as democratic theory suggests.

Politicians are not only grasping to determine what citizens want. They also have a difficult time understanding what citizens need. Rather than being guided by systematic data, politicians are often animated by narratives that focus on how a single individual or a small group of people are helped by a policy. Such empathic personalizing actually impairs good judgment about group needs.¹⁰ When politicians are tested on their knowledge of the material well-being of their constituents, the most recent evidence suggests that they have systematic errors in their perceptions.¹¹

Politicians, for example, do not accurately estimate the financial hardships of their constituents. How can they then be expected to effectively work on citizens' behalf if they do not systematically understand the needs and stations of their citizens?

Finally, politicians (and other public servants) might be expected to care about the effects of the policies which they propose and then implement. But there are limits to this, too. Politicians are often reluctant to pursue information extensively when designing policies¹², and they often show little interest in understanding how well policies are actually affecting citizens on the ground.



CAN WE RETHINK HOW POLITICIANS LEARN ABOUT CITIZENS?

Perhaps it is unreasonable to expect politicians to know everything which a citizen wants them to do. This is especially true given how expensive data has often been. For a long time, learning about citizens' preferences and behaviors was both costly and difficult.¹³ Learning was costly because studies of citizens' preferences often required high quality public opinion studies, which took a long time to collect and were also prohibitively expensive. Learning about their behaviors similarly required active and costly data collection. Moreover, knowing the material status and needs of individual voters with some frequency—i.e. not relying on a decennial census—was out of reach. What is more, learning was often constrained because for ethical or legal reasons, it was not possible to (easily) collect data about citizens.

In many ways, technology changes this relationship. Politicians can get information on citizens' preferences much more easily.

The collection of public opinion data has become remarkably more economical in recent years through various online data methods.

Other methods, like sentiment and text analysis, allows researchers to learn about latent and expressed preferences passively or unobtrusively.

Through advances in “big data” and the processing of high dimensional data, we can also learn more quickly what services citizens are accessing, and how these uses are related to each other. For example, by merging administrative data sources, analysts can learn how access to one government service—for example, income supports—may relate to demand for another service—for example, health care provision. And, through mobility data especially, politicians can learn a lot about patterns of government service usage. In short, the ability to engage data about citizens in policymaking is greater than ever.

HOW COULD POLITICS IMPROVE WITH MORE DATA ENGAGEMENT?

In this section, I provide two sketches of how politicians and public servants could better engage existing high frequency or high dimension data sources to learn more about what citizens want, and what they need and how policies are positively or negatively addressing those needs.

Knowing what citizens want

Knowing what citizens want should be a straightforward enough enterprise. Politicians can simply poll constituents, querying their views on key issues, and then learn from the results of those polls. There are at least four limits to traditional polling techniques, however. First, public opinion studies suffer from substantial problems of non-response bias, where some types of citizens (perhaps a majority) are unwilling to answer surveys, thus limiting the representativeness of survey results. Second, polls are often limited in space, constraining how many issues can be queried or the depth at which they can be sounded. Third, because they can be expensive, polls occur with limited frequency. Finally, polls are typically conducted on samples which, while sufficient to make national level inferences, are nonetheless too small in any single sample to learn about important subgroups, whether demographic or geographic. Consider, for example, a group which represents 5% of the population. In a representative sample of 1000 citizens, just 50 citizens from this group will be present, substantially limiting how much can be learned about that group compared to others.

What can be done in the face of such constraints? There are at least three solutions on offer, which in combination can substantially enhance how much politicians know about citizens preferences. First, more sources than simple polls can be used to measure citizens' preferences. Recognizing that public preferences are often latent—that they lie below the surface but that they are revealed through what people say, how they respond to polls, what they like, forward, and retweet online, even what they buy—we can model how supportive citizens are of some courses of government action over others by correlating large amounts of data across multiple sources.



Second, we can use modern statistical techniques of imputing outcomes—fin this case preferences—to precise demographic groups or geographies, through techniques like multiple regression post stratification. Combining together multiple data sources and then modeling them down to the kinds of small micro-targeted groups politicians care about, can help politicians know what different groups in different geographies want. And by relying on more than just poll data, this can be updated with high regularity.

Imagine, then, a dashboard which for any issue in front of or potentially in front of a legislature would provide a legislator with detailed data about the preferences of citizens, which could be queried at not only a general population level, but for subsets of the population the politician is interested in. Direct, rich, and frequently updated data could empower politicians to represent citizens' preferences much better than they currently do.

Knowing what citizens need

The commercial world is awash in information on consumers. Individuals generate data across thousands of transactions, internet searches, movements, and other behaviors. Importantly, while these are often individual actions, it is possible for these data to be stitched together. Accordingly, we learn not only about

what is happening in the aggregate, but what identified individuals are doing. For any given person, we can potentially understand their movement history, the state of their individual finances, detailed demographic information, information on their personal professional relationships, and even information on their preferences for dating. Many rightly regard this kind of information as intrusive and in violation of basic norms of privacy, an entirely reasonable position. And yet commercial organizations go to great effort to assemble these kind of data within legislated privacy regimes precisely because there is immense value in accurately understanding important information about the lives of consumers.

Do politicians know as much about the people they represent? Do they know how often their constituents are able to access healthy food options within their neighborhoods? Do they know how much individual constituents have to travel for work, commute to receive Medical services, or venture out for recreation? Do they know how often their constituents are expressing concern or experiencing stress over their financial state, through for example search data, accessing their own credit reports, or even asking their financial institutions for short-term help? And, most importantly, can they understand how a policy change would affect any of those things?

The long story short is that in many countries, politicians have a much fuzzier view of the lives of their citizens than the average marketing agency or Swift security corporation, and even less of a sense of how a policy would change individuals' lives. At one end of this extreme is the Canadian case, where federal and provincial bureaucracies do not collect systematic digital health data, where there is almost no easy linkage of data at an individual level across multiple departments, and where the national statistics agency cannot easily access information—even in a highly controlled manner—on individual financial holdings. By contrast, Nordic democracies and Israel collect massive amounts of health and social data on citizens, and incorporate this into policy decision making at a granular level. In place of systematic data on the effects of policies and the lives of their constituents, politicians instead lean on impressionistic accounts from interactions with constituents or from received correspondence, or they turn to information provided by interested lobby groups or policy advisors.

Imagine instead a scenario in which politicians were able to understand at the level of individual constituents, whether their financial situations were improving or deteriorating week over week, whether they were experiencing more or fewer health challenges, or in more general terms, whether they needed more or less intervention from the state to help their lives flourish, and in what areas of their lives. Suppose that politicians could have access to high frequency, high detail data-driven accounts of how well their constituents are faring materially, and what kinds of government assistance could be effectively targeted to them. Imagine too that the tools of causal inference were applied to these data, to understand how the deployment and uptake of support policies actually did or did not improve the lives of citizens. Such policy making would require a substantial bargain between government and citizens: that citizens would be willing to give over large amounts of their data, and that government could be trusted to use these data in a manner consistent with democratic ends and not in a manner that violates privacy or other democratic norms.

THE DEMOCRATIC ADVANTAGE

In the battle for data, surveillance, and learning, it is sometimes argued that autocratic regimes have a natural advantage. They are unconstrained by high citizen expectations of privacy and are often not bound by effective legal regimes which enforce those protections and other rights. Moreover, unconstrained by democratic processes, they can move with great rapidity to change course when they learn about new states of the world. In comparison, democratic government is said to be overly concerned with privacy and handcuffed by partisan fighting and by checks and balances.

This discourse, however, fails to see at least three reasons why democratic governments could benefit from the use of high dimensional data and applied artificial intelligence to learn about and act on citizens preferences in a way that autocratic regimes cannot.

The first reason is the “values premium” that exists within democracies. Embedded within democracies is the notion that how citizens are treated matters as much as what citizens get. Process matters as much as outcomes. Democracies put a premium on values like trust, transparency, and decency. In the evocative example of Avishai Margalit in his work, *The Decent Society*, we are asked to consider the difference between delivering bread in a famine from the back of truck, where in one scenario it is handed to recipients and in another it is thrown at their feet to be scrambled after. In both scenarios the same outcome results: people get food. But only in one are citizens treated decently. That makes all the difference. Democracies are practiced—especially at the level of “street level bureaucrats”—at treating people with decency, as rights-bearing individuals. This premium on values and process and not just on outcomes, is one part of the democratic advantage.

Second, democracies are naturally better at incorporating feedback. This “feedback advantage” comes from the competitive nature of democracies. Autocracies suffer from inefficient feedback mechanisms, as public criticism of state action is regularly short-circuited. Instead of soliciting genuine, organic measures of satisfaction among citizens, autocratic states impose order and assume all is well among citizens. By contrast,

the competitive incentive in democracies to find out what governments have done well and poorly invites constant refinement of processes and policies. The same would apply in cases in which governing officials are using large amounts of data to learn about and act on the preferences of citizens.

Finally, there is a public sector advantage. Compared to private organizations, democratic public services are arguably more culturally ready for the adoption of this technology than any other organization, precisely because public services already resemble human-assisted AI systems. A public servant is already used to working within a prediction machine: they are presented with a problem, they formulate and test solutions using data, and they then make recommendations through a series of considerations—or algorithms—which is eventually placed before a human to make a choice from a small number of options. That final decision maker is the human in the loop, and while they cannot see all the deliberations and considerations that lead to a recommendation, they have a responsibility to own the decision and to be able to explain and justify it if demanded. All these elements map onto a well-designed system of human-assisted AI.

CONCLUSION

Politics is a difficult job, done by humans. Those humans are limited in their capacity to imagine the preferences of others and to understand their needs.

Effective democratic governance depends on us enhancing the capacity of public figures to know and effectively act upon citizens' wants and needs.

The availability of data and our capacity to learn from it is increasing at a breathtaking clip. By taking seriously the insights afforded through the combination of data, artificial intelligence, and machine learning, public servants can better know what it is citizens want them to do and what citizens need them to do, and by leaning in directly on the values already embedded in democratic systems, the need for decency, the need for aligning values with actions, and the need for democratic accountability and explanation, makes our public systems ironically as ready as any to unlock the gains provided by this combination of data and learning tool.



ENDNOTES

1 Fukuyama, F. (1989). The end of history? The National Interest, 16, 3-18.

2 Leonard, M. (2008). What does China think? PublicAffairs.

3 Fox, J., & Shotts, K. W. (2009). Delegates or trustees? A theory of political accountability. The Journal of Politics, 71(4), 1225-1237.

4 Hertel-Fernandez, A., Mildenerberger, M., & Stokes, L. C. (2019). Legislative staff and representation in Congress. American Political Science Review, 113(1), 1-18.

5 Broockman, D. E., & Skovron, C. (2018). Bias in perceptions of public opinion among political elites. American Political Science Review, 112(3), 542-563.

6 Pilet, J.-B., Helfer, L., Sheffer, L., Varone, F., Vliegenthart, R., & Walgrave, S. (Forthcoming). The Conservative Bias Among Politicians: A Five Country Comparative Study. American Political Science Review.

7 Walgrave, S., Jansen, A., Sevenans, J., Soontjens, K., Pilet, J.-B., Brack, N., Varone, F., et al. (2023). Inaccurate Politicians: Elected Representatives' Estimations of Public Opinion in Four Countries. The Journal of Politics, 85(1), 209-222.

8 Kalla, J. L., & Porter, E. (2021). Correcting bias in perceptions of public opinion among American elected officials: results from two field experiments. British Journal of Political Science, 51(4), 1792-1800.

9 Miller, W. E., & Stokes, D. E. (1963). Constituency influence in Congress. American Political Science Review, 57(1), 45-56.

10 Bloom, P. (2017). Against empathy: The case for rational compassion. Random House.

11 Thal, A. (2023). Do Political Elites Have Accurate Perceptions of Social Conditions? British Journal of Political Science. (Forthcoming).

12 Loewen, P. J., Rubenson, D., & McAndrews, J. R. (2022). When Do Politicians Pursue More Policy Information? Journal of Experimental Political Science, 9(2), 216-224.

13 Fenno, R. F. (1977). US House members in their constituencies: An exploration. American Political Science Review, 71(3), 883-917.

FURTHER REFERENCES

Druckman, J. N., & Jacobs, L. R. (2006). Lumpers and splitters: The public opinion information that politicians collect and use. Journal of Public Opinion Quarterly, 70(4), 453-476.

Gilens, M., & Page, B. I. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. Perspectives on politics, 12(3), 564-581.

Henderson, G., Hertel-Fernandez, A., Mildenerberger, M., & Stokes, L. C. (2021). Conducting the heavenly chorus: Constituent contact and provoked petitioning in congress. Perspectives on Politics, 1-18.

