# Problems with a Methodology Used to Evaluate the Voter Initiative

## John G. Matsusaka
### University of Southern California

This note identifies problems with a methodology that has been used to test whether policy is more or less responsive to public opinion in states with voter initiatives. The methodology is to regress a policy variable on a measure of constituent preferences, and compare the coefficients for states with and without voter initiatives. The states with the largest coefficients are said to be more responsive to public opinion. Such an inference is shown to be invalid.

$\mathbf{D}$o voter initiatives make government policies more responsive to the demands of citizens?[1] Several recent studies have attempted to answer this question using different methodologies and have come to different conclusions. Gerber (1996, 1999) finds that initiatives increase policy responsiveness, while Lascher, Hagen, and Rochlin (1996) and Camobreco (1998) find that initiatives decrease (or at least do not increase) responsiveness.[2] The purpose of this note is to highlight a problem in the methodology of the Lascher, Hagen, and Rochlin and the Camobreco papers. The problem is severe: it seems to prevent us from drawing any conclusion from the two papers about whether policy is more or less reflective of public opinion in initiative states.

## The Methodology in Question

The methodology used in Lascher, Hagen, and Rochlin (1996) and Camobreco (1998) is based on a regression of the form,

$$G_i = a + b \cdot P_i + c \cdot X_i + e_i, \tag{1}$$

[1]The voter initiative is a direct democracy institution by which citizens can propose and pass laws directly without the involvement of their elected representatives. California is the most prominent initiative state, but 23 others provide for the initiative, as do roughly half of all cities.

[2]These papers examine different policies. Gerber looks at the death penalty and parental abortion notification; Lascher, Hagen, and Rochlin investigate a menu of policies (including AFDC, "consumer policy," "criminal justice policy," education expenditure, ERA, gambling, Medicaid, and tax progressivity); and Camobreco studies fiscal policies.

where the unit of observation ($i$) is a state, $G$ is some government policy (such as expenditure per capita), $P$ is a measure of public opinion/constituent preferences (such as the state ideology numbers in Erikson, Wright, and McIver (1993), henceforth EWM), $X$ is other control variables, $e$ is an error term, and $a$, $b$, and $c$ are parameters to be estimated. In this regression, $b$ measures the responsiveness of policy to public opinion at the margin. If constituent desires matter at all for policy, we expect $b > 0$.

To evaluate the effect of the initiative on responsiveness, the basic regression in (1) is modified to allow $b$ to be different in initiative and non-initiative states:

$$G_i = a + b_I \cdot I_i \cdot P_i + b_N \cdot N_i \cdot P_i + c \cdot X_i + e_i, \tag{2}$$

where $I$ is a dummy variable equal to 1 for initiative states, and $N$ is a dummy variable equal to 1 for non-initiative states.[3] A test is then performed to compare coefficients. If $b_I < b_N$ the proffered conclusion is that policy in initiative states is less responsive to public opinion than policy in non-initiative states.

## The Problem

While the procedure outlined above has an intuitive appeal, intuition leads us astray. It turns out that a comparison of $b_I$ and $b_N$ can tell us virtually nothing about which institution does a better job satisfying constituent desires.

To see the problem, it is useful to step back from the regressions for a moment and focus on the underlying question: how does the initiative affect the relation between policy outcomes and the desires of citizens? The most direct test of the question would be to compare the actual policy in state $i$, $G_i$, with the policy desired by the public, call it $G_i^*$. If we found that $G_i$ was closer to $G_i^*$ in initiative states than non-initiative states (say, in terms of the sum of squared deviations), then we would conclude that the initiative made government more responsive to citizen demands (and the converse).

Unfortunately, we do not observe $G_i^*$ directly. Instead, we observe a measure of public opinion or preferences, $P$, which is correlated with $G_i^*$. Let us write the relation between preferences and the public's desired policy as $G_i^* = f(P)$, where $f$ is the function that translates opinion into policy. The heart of the problem is that we do not know the form of $f$.

A numerical example will illustrate the problem. Suppose $P_i$ is the EWM ideology score for state $i$, and $G_i$ is the state's per capita government spending. The theoretical range of $P_i$ is $-100$ to $+100$, where negative numbers are "conservative" and positive numbers are "liberal." Spending $G_i$ can take on values

---

[3] Equation (2) is not the specification employed in the papers although it is formally equivalent. The actual specification used is $G_i = a + \beta_1 \cdot I_i \cdot P_i + \beta_N \cdot P_i + c \cdot X_i + e_i$. The transformation is $b_N = \beta_N$ and $b_I = \beta_N + \beta_I$. The problems are easier to see with specification (2).

from 0 to the state's income per capita. Suppose again for the sake of example that the desired level of spending as a function of $P$ is given by

$$G_i^* = 2{,}000 + 20P_i. \tag{3}$$

The $f$ function stated in equation (3) implies that the desired spending level in the most conservative state ($P = -100$) is 0, desired spending in a moderate state ($P = 0$) is 2,000, and desired spending in the most liberal state ($P = +100$) is 4,000.

Now suppose that we regressed observed values of $G$ on $P$ and estimated separate coefficients for initiative and non-initiative states. What conclusion would be justified if we found that $b_I < b_N$? This information in itself probably tells us nothing about which type of state is choosing policies that more closely reflect public opinion. In order to perfectly reflect public opinion, we must have $G = G^*$, and therefore $b = 20$. If the data revealed that $b_I = 20$ and $b_N = 30$, then initiative states would be giving citizens exactly the policy desired while non-initiative states would be spending more than desired. Thus, even though $b_I < b_N$, initiative states are more responsive than non-initiative states. Of course, it could go the other way: if we estimated $b_I = 10$ and $b_N = 20$, then the non-initiative states would be more responsive. *The point is that a finding of $b_I < b_N$ can be consistent with better or worse representation in initiative states, and therefore, the comparison is uninformative about the question of interest.*[4]

In order to assess how representative a state's policies are, we need to compare the empirical $G$-$P$ relation with the $G^*$-$P$ relation, given by $f$. The ultimate problem here is that we do not observe the true form of $f$. Thus, there is no basis for comparison. Figure 1 illustrates this point in a different way. Suppose the data generate a relation between government policy ($G$) and public opinion ($P$) given by $I$ for initiative states and $N$ for non-initiative states. I chose this pattern because it is close to the actual relation for spending shown in Camobreco (1998). Suppose the "desired" policy as a function of $P$ is given by $f_1$. Then, clearly $I$-states come closer to satisfying the public's demands than $N$-states, even though $b_I < b_N$. This is essentially the point made just above.

But suppose that the relation between desired policy and preferences is given by $f_2$. Now the issue of whether initiative or non-initiative states are doing a better job is fundamentally ambiguous. If we use the criterion of absolute deviations from $f_2$, then $I$-states better reflect desired policy when $P < P^*$, while $N$-states do a better job for $P > P^*$. I have constructed this example so that the slope of $f_2$ is the same as the slope of $N$. Thus, $N$-states have the ideal $b$. However, they have the wrong intercept ($a$).

---

[4] EWM (1993, 92–94) make essentially the same point in a different context. The criticism here is analogous to the "multiple fallacy" emphasized by Romer and Rosenthal (1979) in their assessment of tests of the median voter theorem.
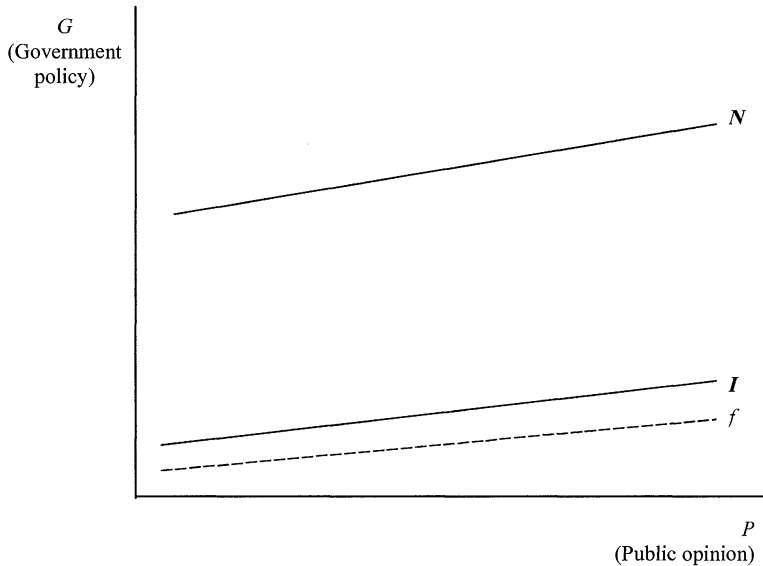
FIGURE 1
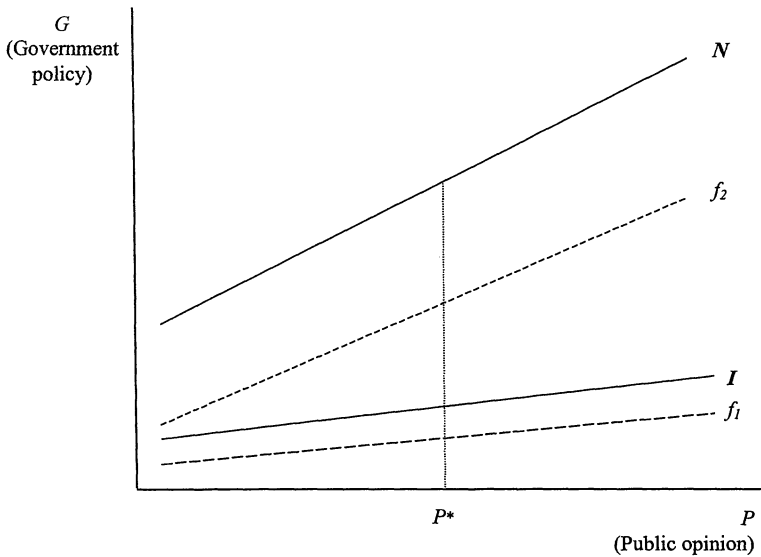


*Note:* $I$ is the empirical relation between policy and public opinion in initiative states, $N$ is the empirical relation in non-initiative states, and $f_1$ and $f_2$ indicate the relation between the public opinion's opinion and its "desired" policy.

By the same reasoning, a finding of $b_I = b_N$ does not imply that initiative and non-initiative states are equally responsive to citizen demands. Consider, for example, Figure 2, where $I$ is the empirical relation between $G$ and $P$ in initiative states, $N$ is the relation in non-initiative states, and $f$ is the "desired" relation. In this case, $b_I = b_N$, but initiative states deliver better policies for their citizens. Thus, failure to find a difference between $b_I$ and $b_N$ (which is the case for some reported regressions) does not lend much support to the hypothesis that initiative and non-initiative states are equally responsive.[5]

Two observations seem appropriate in light of these examples. First, since the function translating public opinion into desired policy is likely to have multiple parameters, we cannot look at a single parameter to determine responsiveness. Other parameters, such as the intercept, should be allowed to vary between initiative and non-initiative states. Second, responsiveness may be fundamentally ambiguous: policies might be more responsive to certain types of public

[5] Another difficulty is that regression (2) would be misspecified if the true relations were those of Figure 2 because it imposes that $N$ and $I$ have the same intercept.

FIGURE 2



*Note:* $I$ is the empirical relation between policy and public opinion in initiative states, $N$ is the empirical relation in non-initiative states, and $f$ indicates the relation between the public's opinion and its "desired" policy.

opinion in initiative states and more responsive to other types of opinion in non-initiative states.[6]

## Conclusion

This note identifies problems with a methodology recently employed to test the responsiveness of policy making in initiative and non-initiative states. The main conclusion is that unless the researcher knows exactly how his or her public opinion variables translate into desired policies, there is no way to determine whether one institution is doing a better job meeting constituent desires than another.[7]

The general issue here—how well do alternative institutions represent constituent interests—is an important one. If the existing methodology is flawed,

[6] Another layer of complexity arises if public opinion is multidimensional. For example, suppose the desired policy depends not only on ideology but also on race, religion, income, and so on. Then, the desired policy $G^*$ is a function of $P_1, P_2, \ldots$, with corresponding coefficients $b_1, b_2, \ldots$

[7] It may be worth noting that the criticism here applies only to *comparisons* of responsiveness between states. If we want to investigate whether policy responds to preferences at all, which was the purpose of EWM (1993), then tests of $b = 0$ in equation (1) are appropriate.

what can be done? One solution is to investigate policies where we can observe directly the desired policy, $G^*$. The natural source of such information is opinion polls that ask about specific policy choices. This information can be used to calculate the difference between actual and desired policies directly. One approach, for example, would be to calculate $\sum_i (G_i^* - G_i)^2$ for initiative and non-initiative states and see which is lower. This is the idea underlying the studies by Gerber (1996, 1999) on death penalty and abortion policies, although her ultimate empirical specification is different because of the way she transforms the model.

If the desired policy cannot be observed directly, it still may be possible to draw some conclusions if (1) the policy space is binary and (2) a fairly direct correlate with preferences is available. An example that fits would be a policy to conduct or prohibit capital punishment and a preference measure indicating "thermometer" survey responses to capital punishment. If the policy space is binary and we have a good measure of preference, it seems reasonable to posit a cutoff preference point above which a person prefers one policy and below which the other policy is preferred. Then $f$ is a one-parameter step function. One approach would be to compare squared deviations in initiative and non-initiative states under alternative assumptions about the cutoff points and show that the results are robust to choices of the cutoff parameter. With some structural assumptions, it also might be possible to recover the cutoff parameter using maximum likelihood techniques. If only indirect preference measures are available, such as a person's liberal-conservative ideology and income, then it will often be necessary to use multiple variables to capture preferences, implying multiple cutoff points, and this approach may become unworkable.

The purpose of this note is to suggest that a methodology used in the literature is flawed, not to take a position in the more important substantive debate that underlies the literature. Still, a few thoughts on where we stand on the substantive issues are probably in order. My view of the literature is that we should be fairly agnostic about whether direct or representative democracy does a better job satisfying public opinion. However, the evidence we do have tends to be favorable for the initiative. Gerber (1996, 1999) studies two policies that may be idiosyncratic—death penalty and abortion—but hers appears to be the only reliable statistical evidence, and it supports the view that policies are closer to what voters want in initiative states. The anecdotal evidence points in the same direction. For example, the evidence in Matsusaka (1995) that initiative states cut spending faster in the 1980s than non-initiative states fits with polling data from the 1970s and 1980s indicating that citizens were growing more conservative.[8] The fact that term limits are virtually coterminous with availability of the initiative is another example. Finally, the theoretical literature is al-

---

[8] The evidence in Matsusaka (2000) that initiative states spent more in the first half of the century also supports this view if voters were eager for more spending during this period. But we need public opinion data to be confident of this interpretation.

most unanimous that initiatives increase responsiveness (see Matsusaka and McCarty (2001) for a review). None of this makes a conclusive case for the benefits of the initiative, but it seems to stack up favorably compared to the case against the initiative.

## References

Camobreco, John F. 1998. "Preferences, Fiscal Policies, and the Initiative Process." *Journal of Politics* 60(August): 819–29.

Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1993. *Statehouse Democracy: Public Opinion and Policy in the American States.* Cambridge: Cambridge University Press.

Gerber, Elisabeth R. 1996. "Legislative Response to the Threat of Popular Initiatives." *American Journal of Political Science* 40(February): 99–128.

Gerber, Elisabeth R. 1999. *The Populist Paradox: Interest Group Influence and the Promise of Direct Legislation.* Princeton, NJ: Princeton University Press.

Lascher, Edward L. Jr., Michael G. Hagen, and Steven A. Rochlin. 1996. "Gun Behind the Door? Ballot Initiatives, State Policies, and Public Opinion." *Journal of Politics* 58(August): 760–75.

Matsusaka, John G. 1995. "Fiscal Effects of the Voter Initiative: Evidence from the Last 30 Years." *Journal of Political Economy* 103(June): 587–623.

Matsusaka, John G. 2000. "Fiscal Effects of the Voter Initiative in the First Half of the Twentieth Century." *Journal of Law and Economics* 43(October): 619–50.

Matsusaka, John G., and Nolan M. McCarty. 2001. "Political Resource Allocation: Benefits and Costs of Voter Initiatives." *Journal of Law, Economics and Organization.* Forthcoming.

Romer, Thomas and Howard Rosenthal. 1979. "The Elusive Median Voter." *Journal of Public Economics* 12(February): 143–70.

John G. Matsusaka is professor of finance and business economics, the Marshall School of Business, University of Southern California, Los Angeles, CA 90089-1427.