The Endogeneity of the Initiative: A Comment on Marschall and Ruhil

John G. Matsusaka, *University of Southern California*

IT IS WIDELY BELIEVED THAT DIRECT DEMOCRACY, in the form of the initiative, brought about cuts in state taxes and spending over the last quarter century. This belief is based on firsthand observation, case studies, and more than a dozen statistical studies (Matsusaka 2004). Marschall and Ruhil (2005) focus on one of the central issues in this literature: Did the initiative cause this perceived reduction in taxes and spending, or was there some unmeasured factor that led to adoption of the initiative as well as spending and tax cuts?

Previous research addressed this problem of spurious correlation in a variety of ways, ranging from detailed examinations of specific cases (for example, Gerber et al. 2001) to explicit attempts to control for missing variables (for example, Matsusaka 1995 and Merrifield 2000). Marschall and Ruhil approach the problem using instrumental variables and arrive at a conclusion diametrically opposite from the previous literature: the initiative increased spending and taxes.

This comment explains why that surprising conclusion is probably mistaken. Marschall and Ruhil are to be applauded for applying careful empirical techniques to an important problem, and their article lays down valuable methodological tracks for other researchers to follow. However, their empirical results are fragile, and their model specification can be rejected on both theoretical and empirical grounds. With a minor change in their specification that can be justified theoretically and empirically, or the inclusion of another instrument, their finding is reversed, coming into conformity with the rest of the literature.

State Politics and Policy Quarterly, Vol. 5, No. 4 (Winter 2005): pp. 356–363

INVALID INSTRUMENTS

Endogeneity problems boil down to a concern about unmeasured variables. If there is an unmeasured variable that affects both policy and whether the initiative is available, estimates of how the initiative affects policy will be biased. Formally, policy Y (spending and taxes) is assumed to be determined by I (whether a state has the initiative), other control variables (X), and an error (u):

$$(1) \quad Y = aI + bX + u$$

Initiative status is itself the result of some other factors (Z) and an error (v):

$$(2) \quad I = cZ + v$$

A problem arises if u and v are correlated. In such an instance, if (1) is estimated alone, then the estimate of b will be biased. Intuitively, if u and v move together, then changes in the errors will be falsely attributed to the effect of I on Y.

Instrumental variables are a standard approach to such problems. Since the formal details are available in any econometrics textbook (for example, Judge et al. 1985), I focus on the intuition and practical implementation. The basic idea is to identify a set of Z variables (the instruments) that account for I, use these instruments to generate predicted values of I that are purged of v by construction, and then use the predicted values in (1) to produce an unbiased estimate of b. In practice, (1) and (2) are estimated simultaneously. In the following, I call the equation corresponding to (2) the first-stage equation and the equation corresponding to (1) the main equation.

When using instrumental variables, everything hinges on the quality of the instruments Z. To be a valid instrument, a variable must be correlated with I and, although sometimes overlooked, it must be uncorrelated with the error term in the main equation. Intuitively, we need Zs that determine whether a state has the initiative process but do not directly influence the policy under investigation. If a variable influences the policy directly, it must be included in the main equation and cannot be used as an instrument.

The critical instrument proposed by Marshall and Ruhil—the size of the state legislature measured by the number of seats—violates the conditions for a valid instrument; there are strong theoretical reasons to believe that the size of the legislature affects spending and taxes directly. Buchanan and Tullock (1962) showed long ago that gerrymandering could cause the policy preference of the median legislator to diverge from the policy preference of

the median voter, and such distortions are potentially larger as the number of seats increases (see also Gilligan and Matsusaka 1995). Weingast, Shepsle, and Johnsen (1981, 645) formalized one version of this idea, calling it "The Law of $1/n$," defined as "the optimum project scale for any district grows as the polity is more finely partitioned into districts." This theory is based on the observation that the beneficiaries of geographically concentrated projects only pay $(1/n)$th of the cost, where n is the number of districts. A growing body of evidence supports this theory (Gilligan and Matsusaka 1995, 2001; Bradbury and Crain 2001; Baqir 2002). Thus, given strong grounds to expect that the size of the legislature influences policy directly, legislature size belongs in the main equation and is not a valid instrument.[1]

This theoretical concern is significant empirically. Table 1 shows what happens to the estimated initiative effect when the legislature size variable is moved to the main equation. To conduct this test, I obtained most of the data used by Marschall and Ruhil (as much as was provided to me). Model 1 essentially replicates the expenditure model in their Table 4 (unreported revenue and tax models tell the same story as the expenditure models in my Table 1). The numbers are not identical to theirs for several reasons. First, Marschall and Ruhil did not provide me with all of their original data so I was forced to construct parts of it from primary sources following the descriptions in their article. Second, I corrected what I believe were some minor errors in their data.[2] Third, I deleted their signature requirement variable. Since this variable is a function of the initiative dummy variable, it is also endogenous and should be instrumented as well. The patterns in Table 1 appear even if the signature variable is included. Otherwise, I follow their specification. Despite these minor modifications, the estimates of Model 1 are fairly similar to the Marschall-Ruhil estimates, and the coefficient on the initiative variable ($117.56) is positive, which is their main finding.

Model 2 in Table 1 includes legislature size in the main equation instead of as instrument (the same picture emerges if it is also included in the instrumental equation). The coefficient for the initiative variable now becomes negative and statistically different from zero (–$125.88). In short, theory implies that legislature size should be in the main equation and, when it is included there, the results revert to the standard finding in the literature.[3]

Furthermore, the estimated value of ρ (the correlation between the errors u and v) is different from zero and statistically significant in Model 1, but approximately zero in Model 2. Once legislature size is included in the main equation, there is no evidence of correlation errors.

Similar concerns could be raised with Marschall and Ruhil's other two instruments. *Professional* legislatures may choose different policies than citi-

Table 1. Alternative Specifications of the Marschall-Ruhil Expenditure Model

	Model 1		Model 2		Model 3	
	Expenditure (Main)	Initiative (1st Stage)	Expenditure (Main)	Initiative (1st Stage)	Expenditure (Main)	Initiative (1st Stage)
Initiative dummy	117.56**	—	−125.88**	—	−84.67***	—
	(50.39)		(59.01)		(25.01)	
Income	0.11***	−0.00	0.11***	0.00	0.11***	0.00***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Population density	−0.24***	—	−0.23***	—	−0.27***	—
	(0.06)		(0.06)		(0.06)	
Metro population	2.66***	−0.02***	1.02	−0.02***	2.15***	0.01
	(0.78)	(0.00)	(0.79)	(0.00)	(0.77)	(0.01)
Population growth	0.14	0.05***	3.40*	0.07***	3.15	0.09***
	(2.04)	(0.01)	(1.97)	(0.01)	(1.90)	(0.03)
Minerals	−0.01	—	−0.01	—	−0.01	—
	(0.01)		(0.01)		(0.02)	
Federal aid	2.05***	—	2.01***	—	2.09***	—
	(0.14)		(0.13)		(0.13)	
Government ideology	0.61	—	0.19	—	0.33	—
	(0.54)		(0.49)		(0.51)	
South	−119.74***	—	−141.32***	—	−125.06***	—
	(26.25)		(25.36)		(27.32)	
Seats	—	−0.01***	−0.96***	—	—	−0.01***
		(0.00)	(0.18)			(0.00)
Professionalism	—	3.65***	—	3.16***	—	−2.52***
		(0.68)		(0.64)		(0.93)
Dissonance	—	0.01**	—	0.01**	—	0.01
		(0.00)		(0.00)		(0.01)
Initiative 1920	—	—	—	—	—	4.42***
						(0.37)
ρ	−0.56***		0.01		0.10	
	(0.12)		(0.15)		(0.12)	

*p < .10; **p < .05; ***p < .01

Note: Variables are defined in Marschall and Ruhil (2005), and the models follow their specification (except for the particular variables included). Three models are reported, each with 441 observations. The dependent variable in each main equation is state government expenditure per capita. The dependent variable in each first-stage equation is a dummy equaling 1 if the initiative is available. Robust standard errors are in parentheses beneath the coefficient estimates. Coefficients on nine year dummies in the main equations and a constant in the first-stage equations are not reported.

zen legislatures, independent of initiative status. Legislatures with ideologies far from that of their constituents (*dissonance*) may also choose policies different from those of legislatures closely aligned with citizens, especially if the main equation controls for citizen preferences. Since these variables plausibly belong in the main equation, they are unsuitable instruments.

Another problem with Marschall and Ruhil's three instrumental variables is their weak theoretical link to initiative availability. Formally, these instruments need not be linked causally to the initiative; they only need to display certain statistical properties. Nevertheless, a good theoretical link provides a check against spurious results arising from data mining. Given enough time and computing resources, it is often possible to find a variable that, if used as an instrument, would cause coefficients to jump around in surprising ways. The requirement to have a compelling theoretical linkage disciplines the search process for instruments.

A final reason to be skeptical of the instruments proposed by Marschall and Ruhil is that initiatives have been used to alter the size, salaries, and procedures of legislatures (see, for example, the list of initiatives in Waters 2003). Thus, it seems likely that the direction of causality runs from the initiative to their instruments rather than the other way around, at least to some extent. Since the instruments are not exogenous, the coefficient estimates are biased in unpredictable ways.

A VALID INSTRUMENT

We have seen there are both theoretical and empirical reasons to doubt the validity of the Marschall-Ruhil instruments and the finding they produce. But the underlying question remains unanswered: What is the effect of the initiative on spending after controlling for endogeneity?

To make progress on this question, we need a valid instrumental variable. Poterba (1995, 181) suggests one: "If there is substantial change over time in voter tastes for government spending, but if institutions . . . are difficult to change, then institutions that were adopted well in advance of the sample period will be 'quasi-experiments' in budget policy." Following Poterba, I will use a state's initiative status in 1920 as an instrument. Nineteen-twenty is a full 40 years before the sample period begins and is separated from the sample period by the Great Depression, a world war, and some extreme demographic shifts. Erikson, Wright, and McIver (1993) find that there is virtually no correlation in state ideology across this period, so we can rule out the possibility that this variable is merely a proxy for state ideology. Initiative status in 1920 satisfies the conditions for a good instrument because it predicts initiative

status in 1960–2000 because of the durability of institutions, yet there is no reason to expect initiative status in 1920 to have a direct influence on fiscal policy 40 years later.

Model 3 in Table 1 adds initiative status in 1920 to the list of instruments. As in Model 2, the estimated initiative effect is negative (–$84.67) and statistically significant. This result emerges even though the dubious legislature size variable is retained as an instrument and not included in the main equation, in effect giving the Marschall-Ruhil instruments the benefit of the doubt. The same pattern appears if the legislature size variable is removed from the first-stage model. Also note that the ρ estimate is again approximately zero and not statistically significant, indicating no evidence of a correlation in errors in Model 3.

The change in the sign of the initiative coefficient when a new instrument is added further undermines the validity of the Model 1 instruments. If those instruments were valid, they should give a consistent estimate of the initiative effect and not change so radically when another instrument is added. Roughly speaking, only the precision of the estimate should change with the addition of a new instrument.[4]

CONCLUSION

While the substantive focus of this comment has been the effect of the initiative on fiscal policy, the issues raised apply more generally to research on institutional effects. I have tried to cast my discussion in the form of constructive strategies for identifying valid instruments to control for endogeneity. To be a valid instrument, a variable should be linked theoretically to the institution in question but, equally important, must not be connected theoretically to the policy that is to be explained. A particularly attractive instrument for an institution is the availability of that institution some time before the sample period. Because institutions are quite durable, past values predict current values. But past values do not directly drive current policy outcomes.

On Marschall and Ruhil's substantive finding, I have identified a number of reasons to be dubious of their conclusion that the initiative increased state spending over several decades. Their empirical finding is remarkably fragile and depends on the use of instruments that are invalid on theoretical grounds. When an invalid instrument is removed or when a less controversial instrument is included (the availability of the initiative in 1920), a statistically significant and robust negative effect of the initiative on spending is found, consistent with numerous other studies. In short, a correctly specified model using instrumental variables gives no reason to discard the conclusion

reached in the rest of the literature that having the initiative reduces spending and taxes.

ENDNOTES

The author would like to thank Oguzhan Ozbas for helpful discussions.

1. Marschall and Ruhil seem to acknowledge this point in their endnote 17, but then brush it aside with the curious justification that Matsusaka (1995) also omits the variable. This is curious since their goal is to correct specification errors in that article. They also cite Crain (1999) as arguing that the configuration, rather than the number of districts, is what influences policy, when in fact Crain's argument is that both matter (1999, 675): "The configuration of districts and not merely the number of districts matter for fiscal performance."

2. Apparently, they applied an incorrect price deflator to the 1995 data and did not take into account the time varying nature of the initiative availability in some states. No material conclusions hinge on these corrections.

3. This finding of a negative coefficient for legislative seats is inconsistent with the Law of $1/n$ as discussed above, but not with gerrymandering theories (for example, Buchanan and Tullock 1962).

4. A formal Hausman specification test comparing Models 1 and 3 fails to reject Model 1 ($\chi^2 = 1.32$). Given the clear signs that Model 1 is mis-specified, the failure to reject is likely a result of too much noise; that is, the Hausman test lacks power.

REFERENCES

Baqir, Reza. 2002. "Districting and Government Overspending." *Journal of Political Economy* 110:1318–54.

Bradbury, John Charles, and W. Mark Crain. 2001. "Legislative Organization and Government Spending: Cross Country Evidence." *Journal of Public Economics* 82:309–25.

Buchanan, James M., and Gordon Tullock. 1962. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor, MI: University of Michigan Press.

Crain, W. Mark. 1999. "Districts, Diversity, and Fiscal Biases: Evidence from the American States." *Journal of Law and Economics* 42:675–98.

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

Gerber, Elisabeth R., Arthur Lupia, Mathew D. McCubbins, and D. Roderick Kiewiet. 2001. *Stealing the Initiative: How State Government Responds to Direct Democracy*. Upper Saddle River, NJ: Prentice-Hall.

Gilligan, Thomas W., and John G. Matsusaka. 1995. "Systematic Deviations from Constituent Interests: The Role of Legislative Structure and Political Parties in the States." *Economic Inquiry* 33:383–401.

Gilligan, Thomas W., and John G. Matsusaka. 2001. "Fiscal Policy, Legislature Size, and Political Parties: Evidence from State and Local Governments in the First Half of the Twentieth Century." *National Tax Journal* 54:57–82.

Judge, George J., W.E. Griffiths, R. Carter Hill, Helmut Lutkepuhl, and Tsoung-Chao Lee. 1985. *The Theory and Practice of Economics*. New York: John Wiley and Sons.

Marschall, Melissa J., and Anirudh V.S. Ruhil. 2005. "Fiscal Effects of the Voter Initiative Reconsidered: Addressing Endogeneity." *State Politics and Policy Quarterly* 5:327–55.

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

Matsusaka, John G. 2004. *For the Many or the Few: The Initiative, Public Policy, and American Democracy*. Chicago: University of Chicago Press.

Merrifield, John. 2000. "State Government Expenditure Determinants and Tax Revenue Determinants Revisited." *Public Choice* 102:25–50.

Poterba, James M. 1995. "Capital Budgets, Borrowing Rules, and State Capital Spending." *Journal of Public Economics* 56:165–87.

Waters, M. Dane. 2003. *Initiative and Referendum Almanac*. Durham, NC: Carolina Academic Press.

Weingast, Barry R., Kenneth A. Shepsle, and Christopher Johnsen. 1981. "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics." *Journal of Political Economy* 89:642–64.