Rev Austrian Econ (2009) 22:159–167
DOI 10.1007/s11138-009-0073-9

Irrational principals

Bryan Caplan

Published online: 5 February 2009

Abstract Timothy Besley's *Principled Agents?* carefully surveys the modern social science literature on political agency problems and tries to chart a sensible middle course between the naive assumption that politicians maximize the public welfare and the pessimism of Virginia-style public choice. However, the literature that Besley showcases is seriously flawed. By building on the empirically discredited rational expectations assumption, it neglects the strongest normative argument *against* political accountability and overlooks the extent to which "agency failures" stem from principal negligence.

Keywords Principal-agent problems · Voter irrationality · Behavioral political economy

JEL codes D72 · D78 · D03

1 Introduction: the principal–agent framework

Timothy Besley's *Principled Agents? The Political Economy of Good Government* (2006) is a gracious effort to find a middle ground between traditional welfare economics and Virginia-style public choice. Traditional welfare economics plays ostrich in the face of government failure, blithely assuming that politicians maximize the well-being of the public. Virginia-style public choice makes the seemingly more realistic assumption that governments are composed of self-seeking individuals who face amazing opportunities to enrich themselves at the public's expense. However, this pessimistic framework ignores the diversity of leaders' motivation and, in any case, fails to explain why the world is not far worse.

To resolve this tension, *Principled Agents?* (henceforth *PA*) gives a guided tour through the modern scholarly literature on principal–agent problems in politics.

B. Caplan (✉)
Department of Economics, Center for Study of Public Choice, and Mercatus Center,
George Mason University, Fairfax, VA 22030, USA
e-mail: bcaplan@gmu.edu

Besley is a major contributor to this body of work. However, his book is primarily a defense of the insightfulness and relevance of the contemporary mainstream approach—an approach he believes has corrected the errors of earlier scholarly debates. Traditional welfare economics erred by implicitly assuming that the well-being of the public is the sole argument in the utility function of every politician. Virginia public choice erred by embracing the opposite assumption of universal venality and by ignoring the ability of political institutions to channel politicians' self-interest in socially beneficial directions. To truly understand democracy, Besley argues that economists must accept two great lessons:

> Good government is in part associated with designing an institutional framework which affects the incentives of those who make policy decisions. But good government is not entirely about incentives: it also requires good leaders—persons of character and wisdom. (Besley 2006: 2)

The rest of the book focuses on formal models that capture these lessons. Besley briefly discusses some stylized facts about democracy and some econometric results from the US states. However, his main task is theoretical: formally modeling if and how voters can make their leaders serve the public interest.

PA is careful, readable, and demonstrates a mastery of the literature. Its sincere effort to reconcile extremes and reach a reasonable middle ground will win over most academic readers. Nevertheless, Besley's book unintentionally showcases the fundamental flaw of modern political economy: the assumption, despite overwhelming evidence to the contrary, that voters have rational expectations—that their beliefs are, on average, true[1]. This means, for example, that if voters *think* that 80% of protectionist measures are socially beneficial, then 80% of these measures *really are* socially beneficial. While the literature freely grants that people are fallible, it insists that their mistakes are random rather than systematic and particular rather than structural.

The rational expectation assumption colors Besley's entire analysis. If, on average, voters understand their own interests, there is no need to second-guess public opinion. The pressing issue for political economists then becomes accountability—how to make leaders do what voters want[2]. In contrast, if voters have deep misconceptions about their own interests, accountability is largely a red

[1] For standard defenses of rational expectations political economy, see Persson and Tabellini (2000), Drazen (2000), Rodrik (1996), Wittman (1995), and Coate and Morris (1995). For my summary of the evidence against voter rationality, see Caplan (2007).

[2] Admittedly, unlike many researchers in this area, Besley recognizes that there is more to good politics than accountability:

> [T]here is no necessary link between accountability and the welfare of society. There is likely to be such a link when the issues at stake are mostly common values… [W]e will work mainly here with models that assume that voters do have a common interest in achieving some outcome and discuss whether we would expect the political system to deliver it. (2006: 101–2)

Notice, however, that Besley's reservations about the benefits of accountability hinge on *differences in values*. But if the rational expectations assumption fails, common values are *not* enough to make accountability a virtue. When voters falsely believe that protectionism will achieve the common value of enriching the country, for example, accountability forces politicians to become protectionists even if they know better.

herring. If the public systematically misunderstands its own interests, who cares whether politicians stray from public opinion?[3] The important topics then become (a) how popular misconceptions distort policy and (b) whether agency failures mitigate these distortions—or amplify them.

The next section of this review briefly reviews the empirical case against the rational voter assumption and explains why blatant voter irrationality is a more compelling explanation for real-world political failure than subtle principal–agent problems. Section 3 goes on to argue that if voters were rational, many real-world political agency failures would be easy to fix. Common-sense institutional changes and voting strategies could put government on a tight leash; political slack persists largely because common sense is not so common. Section 4 concludes.

2 Question your principals: The ubiquity of political irrationality

"At the heart of political agency models," Besley explains, "is the principal–agent relationship between citizens and government; the principals are the citizens/voters while the agents are the politicians/bureaucrats" (2006: 98). Due to imperfect information, politicians do not automatically do what voters desire. Sometimes, they can stray and get away with it: "In the political agency approach, the incentive problem arises because the citizens have delegated authority to policy makers who enjoy an informational advantage" (2006: 99).

Though Besley raises a few caveats, political agency models usually take it for granted that when leaders stray, voters suffer. They take it for granted because they make the standard assumption—so standard that it often goes unstated—that voters have *rational expectations*. Voters may be ignorant about the current "state of the world." However, they have correct beliefs about each state's probability: If π is the objective fraction of honest politicians, we can set citizens' prior probabilities that a politician is honest equal to π without thinking twice—and use their implied posterior probabilities to determine how they vote.

Voters with rational expectations can definitely make decisions that, in hindsight, turn out to be mistakes. As Besley explains:

> Voters may also be poorly informed about the best policy. There are good reasons to think that governments have access to a wide variety of policy advice which should enable them to make better policy decisions than voters left to their own devices. Indeed, the notion of asymmetric information is at the centre of the idea that voters are rationally ignorant (2006: 104)

[3] Besley appears to concede this point at the end of chapter 3:

> It could well be that the role of politicians is to take unpopular actions for the social good even if they do not result in re-election. This is especially true when voters want inconsistent things. Politicians who simply follow voters' interests may actually be worse for society. Of course, the dangers of populism have always been recognized and this is really only an instance of such concerns in an agency framework. (2006: 173)

I will argue, however, that despite Besley's assurances, modern political economy almost completely neglects "the dangers of populism." In standard models, populist inefficiencies only arise if a policy that is bad on average *happens* to be good in a particular case.

But on a deeper level, voters' mistakes make sense. If (a) 50% of politicians are good and the rest are bad, (b) 60% of free trade agreements are bad for voters, and the rest are good, and (c) good politicians only vote for good trade agreements, but bad politicians vote for all of them, then voters who see a leader sign a free trade agreement *should* reduce their P(their leader is good) from 1-in-2 to 2-in-7[4]. If that leader ultimately turns out to be good, it is hardly the voters' fault; they did the best they could with the available information.

Is it not possible, however, that voters suffer from more fundamental errors? The rational expectation assumption is not graven in stone[5]. Perhaps, like many introductory economics students, voters fail to understand the principle of comparative advantage and, therefore, systematically underestimate the social benefits of international trade. If the public misunderstands its own interests in this way, even protectionist policies that *invariably* hurt the majority could easily be popular and democratically sustainable.

There are strong reasons to take concerns about voter bias seriously. For starters, the Psychology and Economics literature offers clear-cut experimental evidence that many systematic biases exist (Rabin 1998; Thaler 1992). Admittedly, this does not prove that bias matters for political agency issues; perhaps, people form better beliefs in the real world of politics than in the artificial world of the laboratory. At minimum, though, the experimental literature shows that political economists ought to be *open to the possibility* that they have built their edifice upon a shaky foundation.

More importantly, a substantial literature has already tested political beliefs for systematic bias and found it to be real and large (Caplan 2007; Althaus 2003). There are three main empirical approaches, all of which reach the same conclusion.

The first is to compare the public's beliefs to objective facts. For example, the National Survey of Public Knowledge of Welfare Reform and the Federal Budget (Kaiser Family Foundation 1995) compares Americans' beliefs about the federal budget to the actual numbers. It finds that the public systematically overestimates the share of the budget devoted to foreign aid and welfare and systematically underestimates the share devoted to Social Security and health. In fact, when asked to name the two largest items in the budget, foreign aid was the *most common* answer, even though it only amounts to about 1%.

The second is the Enlightened Preference approach. One administers a test of objective political knowledge *and* a survey of policy preferences. Then, you statistically estimate if and how greater political knowledge systematically changes policy preferences, controlling for a long list of possible confounding factors. Scott Althaus (2003) thoroughly explores this literature. He finds, contrary to the rational expectations assumption, that—controlling for a long list of potential confounding variables—people who know more want systematically different policies. As a rule, knowledge makes people more economically conservative and more socially liberal. Furthermore, knowledge usually pushes all demographic groups in the same direction. Greater knowledge makes rich and poor alike more pro-market; in

---

[4] P(good leader|trade agreement)$=0.4\times 0.5/(0.4\times 0.5+1\times 0.5)=2/7$.

[5] Besley's book briefly considers this (2006: 172) as an avenue for future research but uncharacteristically provides no citations to existing literature.

fact, knowledge has a substantially *larger* pro-market effect for the poor (Althaus 2003: 111).

The third way to test political beliefs for systematic bias is to compare the beliefs of laymen to those of experts. The presumption is that if laymen and experts systematically disagree, the laymen are probably wrong—especially if these systematic differences persist controlling for many possible confounding factors. This is the primary approach in my book *The Myth of the Rational Voter* (2007). I show that average Americans and PhD economists have large, systematic disagreements about how the economy works, even controlling for income, income growth, job security, race, gender, party identification, ideology, and more[6].

These results should not be surprising to *teachers* of economics. When we teach undergraduates, we know that they arrive as damaged goods. Students think that greedy intentions imply bad social consequences, fail to appreciate the benefits of the international economy, imagine that "saving jobs," not improving productivity, is the key to prosperity, and see an economy going from bad to worse. When we test these stereotypes about non-economists against the data, they all check out (Caplan 2007: 50–93). Only by studying economics do we come to understand the inefficiency of price controls, trade barriers, restrictions on lay-offs and firing, and so on.

Nevertheless, these results should be very surprising to *researchers*. After all, virtually every model in *PA* rules out systematically biased beliefs about economics (or anything else). The mainstream framework does not allow, for example, scenarios where:

- Protectionist policies are socially beneficial with probability .1, but voters *think* such policies are socially beneficial with probability .9.
- Farm programs raise prices by restricting supply, but the majority wants to keep the programs "to make sure there will always be a good supply of food."[7]
- Labor market regulation substantially reduces employment, but the average voter sees no connection.

One of the main motivations behind the political principal–agent literature, of course, is to understand how democracies can sustain inefficient policies like protectionism, farm subsidies, and labor market regulation. If voters know their own interests and politicians require voters' support to get elected, why do special interests come out on top?

PA shows that it is *possible* to answer this question without appealing to voter irrationality. However, this intellectual achievement comes at a high intellectual cost: It implies that *policies that make politicians less popular are, on average, bad for the public*—and that *politicians who adopt policies that are, on average, bad for the public become less popular*. Therefore, when we see a politician support protectionist policies, we can infer either that (a) the politician is hurting his re-election prospects by going against the public's wishes, or (b) for most people, protection is better than free trade. Since it is easy to empirically demonstrate that

[6] Other studies that compare laymen's beliefs to experts' include Kraus et al. (1992) on toxicology, and Lichter and Rothman (1999) on the causes of cancer.

[7] Fifty-eight percent of Americans endorsed this view in a 2004 poll (PIPA-Knowledge Networks 2004).

policies like protection are popular (Poole 2004; Scheve and Slaughter 2001), the mainstream principal–agent literature forces us to conclude that, on average, protectionist policies benefit the majority[8].

Thus, while the principal–agent literature begins by trying to explain inefficient policies, it ends up putting a stamp of approval on the bulk of the status quo. In contrast, once we remove the shackles of the rational expectations assumption, we can finally say the obvious: Democracies tend to adopt inefficient policies because people systematically overestimate the benefits of these policies. Politicians support protection because protection is popular, and protectionism is popular because—as every intro teacher knows—non-economists falsely believe that protection will make most of us richer.

Admittedly, the models in *PA* provide much *subtler* explanations for inefficient policy. But why is subtler better? We know that the public seriously underestimates the social benefits of free trade. We know that free trade is unpopular. Therefore, why do we need a complicated principal–agent model to explain why protectionism happens—especially when that model does violence to the facts?

3 Agency failures as principal negligence

At this point, one could easily object: Even if blatant voter irrationality is the main source of inefficient policy, the principal–agent literature is easy to salvage. We can simply take the models in *PA*, suitably modify the rational expectations assumption, then solve the revised model. This will eventually lead to a nuanced behavioral political economy of the kind that Besley (2006: 172) briefly alludes to.

While this plan sounds reasonable enough, it takes too much for granted. It assumes that agency problems are basically inevitable. We can *cope* with them, but there is no way of *solving* any of them. In this section, however, I argue that political agency problems are often a *byproduct* of voter irrationality. The principals give their agents grossly suboptimal incentives, then complain that the agents fail to carry out their assignments.

For example, a key feature of the main models in *PA* is that there is no pay-for-performance[9]. No matter how good or bad a job a politician does, he gets the same compensation.

Admittedly, this is a standard feature of modern democracies. But *why* is it a standard feature? Because it is too hard to evaluate politicians' job performance? If so, using re-election as a carrot is equally misguided. Because it is too hard to assign optimal weights to various aspects of job performance? If so, one could simply "let the people decide" the optimal weights by basing bonuses on approval ratings[10]. Because

[8] Of course, any particular piece of protectionism could be the exception that proves the rule. But the rational expectation assumption ensures that the rule itself is right.

[9] For a seminal analysis of pay-for-performance in politics, see Walsh (1995).

[10] Finicky rational choice theorists may object that the public would have an ex post incentive to give negative evaluations to all politicians in order to save money. Since individual survey respondents would only have a tiny effect on a politicians' evaluation, however, it is hard to take this concern seriously. Would people really denounce a beloved politician in a phone survey in order to save themselves a fractional penny?

politicians' actions have long-run consequences? If so, bonuses could be a function of long-run consequences. Politicians could continue to earn bonuses long after they have left office (which would also be a good way to handle end-game problems).

Some better arguments against paying politicians for performance may exist (Hart et al. 1997). However, the flimsiness of the leading objections should open us up to a simple alternative: Pay-for-performance is a good idea, but the public is too irrational to accept it. As Caplan (2001: 323) explains:

> Many [voters] prefer to see politicians as altruistic public servants, a breed apart from the self-interested inhabitants of the non-political world. Given public choice scholars' determined efforts to discredit this viewpoint, they can hardly argue that this mistake is not widespread.

Parallel arguments apply to many other so-called agency problems: Simple solutions exist, but the public negligently ignores them. The following are examples.

Politicians have many opportunities to abuse their power, and monitoring is costly This is no reason to despair; as Gary Becker (1968) explained, we can use probability multipliers to make good behavior incentive-compatible despite imperfect information. Indeed, considering how often politicians' scandals and insensitive remarks provoke "public outcry," voters already act as if they understand the Beckerian strategy. So why can they not use the same approach to deter politicians from abusing their power?

Power is too decentralized for Beckerian punishments to work; there is too much "team production" This seems like a reasonable objection, especially in polities like the USA with extensive division of powers and consequent gridlock. However, this just raises another question: If decentralization makes it hard to assign credit and blame, the obvious solution is to centralize. Indeed, many democracies already have parliamentary systems where responsibility is quite transparent; why not copy them?

Beckerian punishments violate the constitutional prohibition against cruel and unusual punishment Efficiency wage models (Shapiro and Stiglitz 1984) point to a simple solution: drastically raise politicians' salaries, then threaten to deprive them of their rents if they misbehave.

Politicians cannot make binding commitments, leading to time inconsistency and other problems Again, this accurately describes modern political reality. However, the obvious solution is to change the law so that politicians *can* make binding commitments. Why not (Landsburg 1993: 148–9)?

My point is not that any of these reforms are likely to happen. Indeed, I expect them to fall on deaf ears. My claim, rather, is that plausible solutions to political agency problems are at hand, but public opinion stands in the way. Political agency problems are probably not a major *independent* source of political failure because they usually require voter irrationality to get off the ground.

4 Conclusion: Salvation through agency "failure"?

PA is a cut above most of the literature it surveys. Researchers in this area tend to uncritically embrace the goal of making politicians do whatever the voters desire. Besley admits more than once that maximum political accountability may not be socially optimal.

Nevertheless, the underlying theme of *PA* is that the mainstream approach is basically sound. Despite a few caveats, this approach takes the rational expectation assumption for granted. As a result, Besley's book reflects researchers' strange conclusion that the teaching of economics serves little social function. In the mainstream principal–agent framework, economists cannot lament the folly of protection or farm subsidies; all they can do is explain why the public's strategies are optimal responses to average conditions.

Casual acceptance of the rational expectation assumption also probably explains why even Besley treats perverse political incentives as exogenous facts of nature, instead of endogenous responses to public opinion. If voters tolerate a system where "dissonant" politicians can abuse their power with little cost, complaints about "agency problems" are misleading. The principals are only reaping what they have sown. They should have known better.

If I am right about public opinion, the great mystery of politics is "Why aren't policies a lot worse?" The answer, almost by definition, is "agency failure": Politicians fail to deliver the specific policies that the public wants. The upshot is that agency failure is the cause of a great deal of *success*. Like monopoly in an industry with negative externalities, agency failures help to mitigate the damage caused by irrational expectations about policy.

Thus, my claim is not that agency problems are uninteresting. My complaint, rather, is that the vast literature surveyed in *PA* downplays the most interesting questions. Where do agency problems come from? What policies would democracies actually supply in the absence of agency problems? How precisely would these policies be inefficient? How do agency problems make these inefficiencies worse? How do agency problems make them better? Before political economists can answer these questions, however, they have to put aside their assumptions about how their voter/principals have to be and take a good hard look at how they actually are.

Acknowledgment I thank Diana Weinert for the excellent research assistance and the Mercatus Center for financial support.

References

Althaus, S. (2003). *Collective preferences in democratic politics: Opinion surveys and the will of the people*. Cambridge: Cambridge University Press.

Becker, G. (1968). Crime and punishment: an economic approach. *Journal of Political Economy*, *76*(2), 169–217.

Besley, T. (2006). *Principled agents? The political economy of good government*. Oxford: Oxford University Press.

Caplan, B. (2001). Rational irrationality and the microfoundations of political failure. *Public Choice*, *107* (3/4), 311–31.

Caplan, B. (2007). *The myth of the rational voter: Why democracies choose bad policies*. Princeton, NJ: Princeton University Press.
Coate, S., & Morris, S. (1995). On the form of transfers to special interests. *Journal of Political Economy*, *103*(6), 1210–35.
Drazen, A. (2000). *Political economy in macroeconomics*. Princeton, NJ: Princeton University Press.
Hart, O., Shleifer, A., & Vishny, R. (1997). The proper scope of government: theory and application to prisons. *Quarterly Journal of Economics*, *112*(4), 1127–1161.
Kraus, N., Malmfors, T., & Slovic, P. (1992). Intuitive toxicology: expert and lay judgments of chemical risks. *Risk Analysis*, *12*(2), 215–232.
Landsburg, S. (1993). *The armchair economist: Economics and everyday life*. NY: The Free Press.
Lichter, S. R., & Rothman, S. (1999). *Environmental cancer—A political disease?*. New Haven, CT: Yale University Press.
Persson, T., & Tabellini, G. (2000). *Political economics: Explaining economic policy*. Cambridge: MIT Press.
PIPA-Knowledge Networks Poll. (2004). "Americans on Farm Subsidies." Retrieved from http://www.pipa.org/OnlineReports/FarmSubsidies/FarmSubs_Jan04/FarmSubs_Jan04_quaire.pdf.
Poole, W. (2004). Free trade: why are economists and noneconomists so far apart? *Federal Reserve Bank of St. Louis Review*, *6*(5), 1–6.
Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, *36*(1), 11–46.
Rodrik, D. (1996). Understanding economic policy reform. *Journal of Economic Literature*, *34*(1), 9–41.
Scheve, K., & Slaughter, M. (2001). *Globalization and the perceptions of American workers*. Washington, DC: Institute for International Economics.
Shapiro, C., & Stiglitz, J. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, *74*(3), 433–444.
Thaler, R. (1992). *The winner's curse: Paradoxes and anomalies of economic life*. Princeton, NJ: Princeton University Press.
Walsh, C. (1995). Optimal contracts for central bankers. *American Economic Review*, *85*(1), 150–67.
Wittman, D. (1995). *The myth of democratic failure: Why political institutions are efficient*. Chicago: University of Chicago Press.