Fundamental Errors in the Voting Booth*

Edward L. Glaeser
Harvard University

Giacomo A. M. Ponzetto
CREI, Universitat Pompeu Fabra, IPEG and Barcelona GSE

August 2017

Abstract

Psychologists have long documented that we over-attribute people's actions to innate characteristics, rather than to luck or circumstances. Similarly, economists have found that both politicians and businessmen are rewarded for luck. In this paper, we introduce this "Fundamental Attribution Error" into two benchmark political economy models. In both models, voter irrationality can improve politicians' behavior, because voters attribute good behavior to fixed attributes that merit reelection. This upside of irrationality is countered by suboptimal leader selection, including electing leaders who emphasize objectives that are beyond their control. The error has particularly adverse consequences for institutional choice, where it generates too little demand for a free press, too much demand for dictatorship, and responding to endemic corruption by electing new supposedly honest leaders, instead of investing in institutional reform.

Keywords: Fundamental Attribution Error, Political Economy
JEL codes: D72, E03

1 Introduction

Can limited rationality explain why voters sometimes accept dictators who subvert democracy, as they did in Nazi Germany? Can it explain why voters are sometimes willing to accept government domination of the free press or limits on free speech? Can semi-rationality explain why voters respond to endemic corruption simply by replacing one set of politicians with another, instead of demanding institutional reform?

*E-mail: eglaeser@harvard.edu, gponzetto@crei.cat. We are grateful to Eduard Llorens for research assistance. Ponzetto acknowledges financial support from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant 714905), the Spanish Ministry of Economy and Competitiveness (grants RYC-2013-13838 and SEV-2015-0563), and the Government of Catalonia under the CERCA program.

A robust psychological literature affirms a tendency, sometimes termed the Fundamental Attribution Error, to attribute people's actions and their outcomes to the agents' innate characteristics rather than to luck or circumstances.[1] Economics research documenting that governors (Wolfers 2007) and CEOs (Bertrand and Mullainathan 2001) are rewarded for luck similarly suggests that observers confuse exogenous shocks for true leadership qualities. In this paper, we examine how two benchmark political models change if voters overestimate the importance of politicians' characteristics, and underestimate the role of incentives and chance. We find that the Fundamental Attribution Error can improve the incentives of incumbents, harm the selection of leaders, and generate too much demand for dictatorship, too little demand for transparency, and fighting corruption by throwing the incumbent rascals out rather than by reforming political incentives.

Following a discussion of the Fundamental Attribution Error (or FAE) in Section 2, Section 3 adjusts the standard signal-jamming model of Alesina and Tabellini (2008) to allow for the FAE. In this model, the provision of public services results from a combination of the politician's ability, luck and expenditure (or effort). Politicians spend on public services, or exert more effort, to make voters think that they are more able. Voters are not fooled, but in equilibrium, politicians still exert effort or spend to shape voter beliefs.

In this context, the FAE implies that voters overstate the variance of politicians' ability and understate the variance of their luck. Consequently, they believe that the signal-to-noise ratio in political outcomes is higher than it is in reality. Without the FAE, politicians ignore services which are largely determined by luck, but this tendency is reduced by the FAE. In many cases, the FAE can actually improve the behavior of politicians overall, because it causes voters to respond more sharply to good or bad outcomes, which they credit excessively to the politician's innate ability. This positive side product of voter error closely follows Ashworth and Bueno de Mesquita (2014), who emphasize that irrational voting may improve politician's incentives in many settings.

Against this benefit, there is a cost. Voters re-elect the wrong people, and choose the lucky over the competent. They focus too much on areas where politicians have little impact. A real-world parallel of this error may be that US voters often seem to elect Presidents to manage the economy rather than to deftly handle global affairs, despite the fact that presidential control over war and diplomacy is arguably far greater than presidential control over GDP growth.

But while the core model presents an ambiguous trade-off between better political behavior and worse political selection, the FAE is unambiguously bad for institutional design. We model the demand for the free press by assuming that for a cost, voters are able to improve their ability to distinguish signal from noise. The FAE means that voters already think that they know much of the true signal. Consequently, they are less interested in external sources of information. The overconfidence created by the FAE may explain why most voters in some countries today seem relatively unfazed by state control over the press.

In Section 4, we turn to a benchmark model of political signaling, on the lines of Besley (2007). In this model, politicians differ in their taste for corruption. Some are always honest,

[1] We briefly review the long literature on the Fundamental Attribution Error in Section 2, and note some of its many nuances. We focus on a particularly simple variant of the error, but it seems the most empirically defensible and the smallest deviation from rationality needed for our modeling approach.

some are never honest, and some are opportunists who cheat only when the price is right. Politicians serve two terms, and in the first term, some opportunists refrain from corruption in order to pool with the honest politicians and improve their probability for re-election.

The FAE in this context means that voters underestimate the share of opportunists, and overestimate the share of always good and always bad leaders. Once again, this error tends to improve the behavior of first term politicians. Since voters incorrectly over-attribute good behavior to good character, more opportunists behave well in their first term. This benefit is offset by a tendency to reelect too many politicians who behave well and too few who behave poorly.

Once again, the larger cost of the FAE comes when we allow voters to select institutions. In this case, we allow voters to enable a first-term leader to become the leader in perpetuity, essentially converting a democracy into a dictatorship. When the first-term leader has behaved well, voters think that he will always behave well, even if the incentives created by reelection disappear. Consequently, the FAE will cause voters to be far too enthusiastic about replacing democracy with dictatorship.

This model also predicts that FAE voters would have little interest in institutional reform rather than just replacing the current leader. At the extreme, if voters believe that all leaders are either good or bad, then incentives are useless. Consequently, the FAE produces the focus on personality rather than institutional structure that is again arguably a feature of many of the world's more corrupt democracies.

In Section 5, we discuss the normative and positive implications of these two models. In Section 6, we broaden the discussion of the FAE to discuss voting on policies. While this section is tentative, it suggests that voters may underestimate the ability of taxes to distort behavior. This can create too much redistributive taxation and too little taxation to curb externalities. Similarly, they may underestimate the tendency of new infrastructure or free parking to elicit a behavioral response from their neighbors, and this may make these policies overly popular.

Section 7 concludes. We do not claim that the FAE is ubiquitous, but it seems plausible that many people—including ourselves—occasionally suffer from it, and Wolfers (2007) shows that politicians are rewarded for luck. If the FAE does operate in the voting booth, then in normal circumstances this has both benefits and costs. The incentives of politicians can often be enhanced by this error, while the selection of politicians is generally harmed.

The larger problem associated with attributing too much to an individual rather than their incentives is that voters will have little interest in changing those incentives. Consequently, they will have too little interest in investing in institutional reform to fix corruption and too much interest in allowing dictators to operate without any democratic incentives whatsoever.

2 The Fundamental Attribution Error

The fundamental attribution error or FAE was named by Ross (1977) who defined it as a "general tendency to overestimate the importance of personal or dispositional factors relative to environmental influences." The psychological literature contains many nuances around this idea, but we will treat the FAE solely as a tendency to attribute outcomes to fixed personal characteristics, rather than transient elements related to incentives, information or other temporary conditions.[2] Voters who experience economic success will tend to give too much credit to their political leaders. Voters who observe good behavior from a politician will tend to think that the politician is inherently good, not that the politician feared punishment for misbehavior.

We model the FAE by assuming that if an outcome reflects a combination of an enduring personal trait and temporary external factors, then voters over-estimate the variance of the personal trait relative to the variance of external factors. In our signal jamming model, based on Persson and Tabellini (2000) and Alesina and Tabellini (2007, 2008), voters attempt to deduce the politician's ability based on an outcome that reflects fixed ability, luck and effort. The FAE then means that voters over-estimate the variance of ability relative to luck. In our signaling model, based on Besley (2007) and Rogoff (1990), voters attempt to screen permanently honest politicians from politicians who are permanently dishonest or opportunistic. In that context, the FAE means that voters underestimate the share of politicians who opportunistically respond to incentives, which is equivalent to overestimating the heterogeneity of individual preferences.

2.1 Evidence on the Fundamental Attribution Error

Our first application of the FAE is most closely tied to Wolfers (2007). He examines gubernatorial elections, and finds that governors who receive lucky breaks, such as governors of oil producing states in periods when the global oil price increases, are more likely to be re-elected. This work echoes Fair's (1978) classic finding that economic conditions have a powerful impact on election results. For economists who believe that politicians have limited power over the economy, these results also support the existence of the FAE. Bertrand and Mullainathan (2001) show the CEOs are also rewarded for luck, which seems to support the power of the FAE in the boardroom. All of these papers suggest that observers are underestimating the importance of luck in driving outcomes.

Our second application of the FAE is more closely tied to the work in behavioral economics emphasizing that people attribute the behavior of others to their innate type rather than incentives or local conditions. Eyster and Rabin (2005), for example, explain the winner's curse by assuming that players don't understand how the actions of others are shaped by their information, which in turn determines their incentives. Dal Bó, Dal Bó and Eyster (2017) show that game participants "systematically underappreciate the extent to which policy changes" that change incentives "will affect the behavior of other people." Camerer, Ho and Chung (2004) argue that many experimental outcomes can be explained if some individuals believe that the behavior of others is essentially random and unrelated to incentives.

The view that observers underestimate the impact of incentives is closely tied to Jones and Harris (1967), a seminal paper in the FAE-related literature. Jones and Harris (1967)

[2] Ross was drawing on older research, especially Jones and Harris (1967) and Jones and Nisbett (1970). Jones and Nisbett had articulated a similar concept "actor-observer divergence," in which actors focus on the role of "external conditions" in driving their own behavior, while observers emphasize "stable dispositional properties of the actor." Later Gilbert and Malone (1995) would define a third related concept "correspondence bias" which is defined as "the tendency to draw inferences about a person's unique and enduring dispositions from behaviors that can be entirely explained by the situations in which they occur."

assigned Duke undergraduates to write a 200 word essay either supporting or attacking the Cuban regime of Fidel Castro, taking a stance that was either chosen or assigned by lottery. Subsequently, other students were asked how much the writer actually supported Castro, using a 10 to 70 point scale.

When the essay's slant was voluntary, the observers assigned a score of 17 to essay writers who wrote against Castro and a score of 60 to those who supported Castro. More surprisingly, when the slant was involuntary, the observers gave a score of 23 to those who wrote against Castro and 44 to those who supported him. While it is possible that the observers just didn't believe that the essay's slant was randomly assigned, Jones and Harris (1967) do perform a number of interventions to try to make the random assignment particularly salient.

Jones and Nisbett (1971) particularly cite MacArthur (1970), who also supports the view that observers believe others are motivated by intrinsic tastes. This study enlists students in a survey about personal relationships and then asks observers why the students were willing to participate in the survey. While the participants themselves cited their particular interest in the survey, the observers emphasized that the participants had "a personal inclination to take part in surveys."

Bierbrauer (1974) followed up Milgram's (1963) classic experiment in which students are told to administer electric shocks to accomplices of the experimenter. Bierbrauer's subjects predicted far too much disobedience in the population as a whole, thereby missing the power of the authority figure in this particular setting. Moreover, since they had observed one person administering the shock, they "assumed that the particular subject's obedience reflected his distinguishing personal dispositions."

Many studies have subsequently investigated the Fundamental Attribution Error, and not all results have affirmed the bias. Malle's (2006) meta-analysis finds a relatively weak overall distinction between actors and observers, but this does not disprove that both actors and observers overstate the importance of disposition relative to situation. Despite remaining empirical uncertainties, and buoyed by the real world political results of Wolfers (2007) and others, we now proceed to explore the political impact of a tendency to over-attribute outcomes to permanent personal characteristics of leaders.

3 The FAE in a Signaling-Jamming Model of Political Agency

In this section, we consider a politician who allocates resources across public services and to himself. The politician's budget is fixed, and he balances the benefits of rent extraction today against the costs of a reduced probability of election, as in Alesina and Tabellini (2008). Voter welfare depends both on the level of rent extraction and on the match between their preferences and the politician's allocation of resources across public services.

3.1 Preferences and Technology

There is a unit mass of voters who derive instantaneous utility from public services according to the utility function:

$$u_t = \sum_{g=1}^{G} \alpha_g \ln y_{g,t}, \tag{1}$$

where $y_{g,t} \geq 0$ denotes the provision of public service g and $\alpha_g \geq 0$ its importance for voter welfare. Since the overall government is fixed, we do not explicitly consider the cost of tax payments in our measure of voter welfare.

Each public service is produced with the technology:

$$y_{g,t} = e^{\eta^{i_t}_{g,t-1} + \eta^{i_t}_{g,t} + \varepsilon_{g,t}} x^{\rho_g}_{g,t}, \tag{2}$$

where $x_{g,t} \geq 0$ denotes expenditure on inputs for service g, $\rho_g \geq 0$ denotes returns to spending, and productivity reflects both exogenous economic conditions $\varepsilon_{g,t}$ and the ability of the ruling politician. Politicians' ability follows a first-order moving average process: in each period t, it is the sum of an inherited component $\eta^{i_t}_{g,t-1}$ and an innovation $\eta^{i_t}_{g,t}$. An intuitive interpretation of this structure is that the government is run by a ruling party that contains overlapping generations of politicians. The productivity of public-service provision then reflects a mix of abilities. The contemporaneous shock $\eta^{i_t}_{g,t}$ reflects the ability of untested junior politicians who will lead the party in the next election. The lagged shock $\eta^{i_t}_{g,t-1}$ is the ability of senior politicians who led the party to victory in the previous election, but will retire at the end of their term in office.

All shocks are independent of each other and across services, periods and politicians. They are jointly normally distributed with mean zero. The variance of the period-t shock to productivity in the provision of service g equals σ^2_g. A share ν_g of this variance is due to exogenous conditions and $1-\nu_g$ to the innovation in politicians' ability. Formally, exogenous conditions in period t have a distribution $\varepsilon_{g,t} \sim N\left(0, \nu_g \sigma^2_g\right)$, while the innovation in the ruling politician's ability has independent distribution $\eta^{i_t}_{g,t} \sim N\left(0, (1-\nu_g)\,\sigma^2_g\right)$.

The ruling politician in period t allocates an exogenous budget $b > 0$ to the inputs required for providing the various public services, but also to socially unproductive expenditures that provide him with rents. His objective is to maximize the present value of the rents he can extract while in office, discounted by the discount factor $\delta \in [0,1]$. For any period t while he is in office, rent extraction equals

$$r_t = b - \sum_{g=1}^{G} x_{g,t}. \tag{3}$$

Voter welfare is separable in expenditure, ability, and exogenous shocks:

$$u_t = \sum_{g=1}^{G} \alpha_g \left(\eta^{i_t}_{g,t-1} + \eta^{i_t}_{g,t} + \varepsilon_{g,t} + \rho_g \ln x_{g,t}\right). \tag{4}$$

As a consequence, the welfare-maximizing budget allocation is time-invariant and independent of ability and exogenous shocks:

$$x^*_{g,t} = \frac{\alpha_g \rho_g}{\sum_{j=1}^{G} \alpha_j \rho_j} b \text{ and } r_t = 0. \tag{5}$$

By definition, voters dislike rent extraction. Intuitively, they desire more expenditure on services that have a greater impact on their utility (α_g) and whose production exhibits greater returns to spending (ρ_g).

3.2 Electoral Discipline

Politicians are incentivized to provide public goods because they face voters at the end of each period, and if dismissed they will never return to power. They are unable to make policy commitments, so their re-election depends on voters' evaluation of their track record, following a classic model of political career concerns (Alesina and Tabellini 2008).

Events in each period t unfold according to the following timeline.

1. The inherited component of the incumbent's ability $\eta_{g,t-1}^{i_t}$ is publicly revealed
2. The incumbent allocates the budget to inputs $x_{g,t}$ and rent r_t.
3. The novel component of the incumbent's ability $\eta_{g,t}^{i_t}$ and economic conditions $\varepsilon_{g,t}$ are realized, but not directly observed by voters. They determine the provision of public services $y_{g,t}$, which is publicly observed.
4. An election is held, pitting the incumbent against a random challenger drawn from the same pool.

In stage 2, when the politician chooses productive expenditure and rent extraction, he has no private information. He is as uncertain as the voters about his ability and exogenous conditions. Moreover, such uncertainty is identical whether the politician has previously won re-election or is in his first term in office. As a result, there is a stationary rational expectations equilibrium in which voters correctly anticipate that every politician in every period chooses the same investment $\bar{x}_g$ in public services and extracts invariant rent $\bar{r}$.

Voters anticipate that the future budget allocation is independent of the identity of the ruling politician. Moreover, they understand that politicians' past abilities $\eta_{g,t-1}^{i_t}$ are impermanent and will not matter in the future. Voters then rationally re-elect the incumbent if and only if they perceive him as more capable than average: formally, if and only if

$$\sum_{g=1}^{G} \alpha_g \tilde{\mathbb{E}} \left(\eta_{g,t}^{i_t} | \eta_{g,t-1}^{i_t}, x_{g,t} \right) \geq 0, \tag{6}$$

where $\tilde{\mathbb{E}}$ denotes the biased expectation of voters subject to the FAE.

We model the FAE by assuming that voters know the true variance of performance σ_g^2, but they misperceive the share that is due to economic conditions: instead of having the correct prior ν_g, they have a biased prior $(1-\beta)\,\nu_g$, where $\beta \in [0,1]$ measures their psychological bias. This assumption means that voters overstate the importance of the person relative to the situation, as in the classic experiments on the FAE. This bias is also highly compatible with the empirical findings of Wolfers (2007). As a result of their error, voters infer the incumbent's ability:

$$\tilde{\mathbb{E}} \left(\eta_{g,t}^{i_t} | \eta_{g,t-1}^{i_t}, y_{g,t} \right) = \left[1 - (1-\beta)\,\nu_g\right] \left(\ln y_{g,t} - \rho_g \ln \bar{x}_g - \eta_{g,t-1}^{i_t} \right). \tag{7}$$

Voters' mistake lies in believing that the variance of politicians' ability $\eta^i_{g,t}$ is $\tilde{\sigma}^2_g \equiv [1 - (1-\beta)\,\nu_g]\sigma^2_g$. The tendency to exaggerate the variability of individual characteristics across people is a hallmark of the FAE. Since voters believe with certainty in their erroneous prior parameter, they never revise it based on the history of realized ability innovations.

The politician perfectly understands the voters' bias. It is immaterial whether he shares it, or holds unbiased priors, or is subject to the FAE to a different extent than his voters. All that matters for electoral incentives is that he should have correct second-order beliefs about voters' assessment. Then he knows he is going to be re-elected if and only if

$$\sum_{g=1}^{G} \alpha_g \left[1 - (1-\beta)\,\nu_g\right] \left[\eta^{i_t}_{g,t} + \varepsilon_{g,t} + \rho_g \left(\ln x_{g,t} - \ln \bar{x}_g\right)\right] \geq 0. \tag{8}$$

Crucially, the incumbent's incentives do not depend on the true impact of ability on performance $(1 - \nu_g)$, but exclusively on the voters' biased perception of this impact $(1 - (1-\beta)\,\nu_g)$. This explains why politicians' first-order beliefs about the relative variance of ability and economic conditions do not impact their behavior. More important, it implies that biased voters, albeit unaware of their bias, have rational expectations over the budget allocation: $x_{g,t} = \bar{x}_g$ for all t. On the equilibrium path, rational expectations imply that the incumbent wins each election with 50% probability.

As we prove in the appendix, the unique stationary rational expectations equilibrium has the following characterization.

Proposition 1 *The ruling politician extracts rents*

$$r = b\left[1 + \sqrt{\frac{2}{\pi}}\frac{\delta}{2-\delta}\frac{\sum_{g=1}^{G}\alpha_g\left[1-(1-\beta)\,\nu_g\right]\rho_g}{\sqrt{\sum_{g=1}^{G}\left\{\alpha_g\left[1-(1-\beta)\,\nu_g\right]\sigma_g\right\}^2}}\right]^{-1}$$

and allocates the remainder of the budget to public services so that

$$\frac{x_g}{x_j} = \frac{\alpha_g \rho_g \left[1-(1-\beta)\,\nu_g\right]}{\alpha_j \rho_j \left[1-(1-\beta)\,\nu_j\right]}.$$

He is re-elected if and only if

$$\sum_{g=1}^{G} \alpha_g \left[1-(1-\beta)\,\nu_g\right]\left(\eta^{i_t}_{g,t} + \varepsilon_{g,t}\right) \geq 0.$$

The FAE does not preclude rational expectations and does not create an incumbency advantage or disadvantage. At the same time, it changes political career concerns because it redirects voters' attention across public services. This implies, first, that politicians are screened along different dimensions of their ability. A second consequence is that incumbents react by devoting resources to different categories of productive public spending. The third and final outcome is a change in the equilibrium amount of rent extraction in response to changes in incentives. In the following section, we characterize these three distortions and their impact on voter welfare.

We begin here by highlighting an immediate implication of Proposition 1: the FAE has no effect on political agency if the government provides a single public service, or if all public services reflect in identical proportions exogenous conditions and politicians' abilities.

Corollary 1 *Suppose all public services reflect in identical proportions exogenous conditions and politicians' abilities ($\nu_g = \nu$ for all g). Regardless of voter bias, the ruling politician extract rents*

$$r = b\left[1+\sqrt{\frac{2}{\pi}}\frac{\delta}{2-\delta}\frac{\sum_{g=1}^{G}\alpha_g\rho_g}{\sqrt{\sum_{g=1}^{G}\left(\alpha_g\sigma_g\right)^2}}\right]^{-1}.$$

and allocates the remainder of the budget optimally across public services. He is re-elected if and only if

$$\sum_{g=1}^{G}\alpha_g\left(\eta_{g,t}^{i_t}+\varepsilon_{g,t}\right)\geq 0.$$

If all public services are equally informative about the incumbent's ability, both his screening and his allocation of productive public expenditure across different public services are optimal irrespective of voter bias. As a consequence, politicians' incentives are also independent of the FAE.

Rent extraction consumes a constant proportion of the budget and reflects the fundamental forces of career concerns. Rents are lower when politicians are more patient ($\partial r/\partial\delta < 0$) because greater patience increases the incumbent's willingness to refrain from current rent extraction in order to gain re-election and extract rents in the future. Rents are lower when returns to government spending are higher ($\partial r/\partial\rho_g < 0$) because higher returns make spending a more effective instrument of signal jamming: by sacrificing the same amount of rent, the incumbent can fake a greater amount of ability. Rents are higher when public-service provision is more volatile ($\partial r/\partial\sigma_g^2 > 0$) because greater volatility makes signal jamming less effective: no matter how much the incumbent tries to surprise the voters with productive public spending, the election is going to be decided instead by large swings in his ability and in exogenous circumstances (Alesina and Tabellini 2007).

The balance of the latter two effects explains why the FAE influences electoral discipline exclusively through differences across public goods. If there is a single public good, voter bias induces voters to over-infer ability from public spending. As a result, the incumbent is incentivized to extract lower rents. However, voter bias identically induces voters to over-infer ability from random shocks that politicians cannot control. As a consequence, the incumbent is incentivized to extract higher rents. The effects of these two over-inferences are perfectly offsetting, so incentives are independent of voter bias.

This independence hinges on the absence of any other voter bias. We have assumed that voters screen politicians optimally—conditional on their imperfect information—when they are not subject to the FAE. Corollary 1 then shows that screening remains optimal if voters are subject to the FAE when all public services are identically informative about the incumbent's ability.

We could have assumed instead that voters are also impressionable, in the terminology of Grossman and Helpman (2001). This additional bias would make them imperfect screeners because they would cast their ballot based not only on their inference of ability, but also on the perceived likability of the candidate, which would swing vote despite being truly orthogonal to subsequent welfare. This probabilistic voting assumption would introduce an intensive margin of electoral support, so the probability of re-election would increase smoothly with voters' inference of the incumbent's ability rather than jumping from zero to one when the posterior crosses zero (Boffa, Piolatto and Ponzetto 2016). As a consequence, with a single public service the FAE would raise welfare by alleviating the mistakes of impressionable voters: over-inference of ability from observed outcomes would lead them to pay less attention to candidates' likability, improving both screening and incentives and thus reducing rent extraction.

3.3 Comparative Statics

Voters' re-election decision is first of all an attempt to select the best politicians. The FAE unambiguously makes such screening less effective.

Proposition 2 *The FAE raises the government's average equilibrium ability at providing a public service if and only if its provision depends relatively more on exogenous conditions and relatively less on government ability than the provision of other public services:*

$$\frac{\partial \mathbb{E}\eta_{g,t-1}^{i_t}}{\partial \beta} > 0 \text{ if and only if } \sum_{j=1}^{G} (\nu_g - \nu_j)\, \alpha_j^2 \tilde{\sigma}_j^2 > 0.$$

This redirection of screening reduces voters' welfare ($\sum_{g=1}^{G} \alpha_g \partial \mathbb{E}\eta_{g,t-1}^{i_t} / \partial \beta \leq 0$).

The FAE redirects voters' attention from skills that are truly reflected in government performance to others that biased voters incorrectly infer on the basis of exogenous circumstances the government cannot actually control. As a result, voters select politicians who are more skilled at providing public services whose provision has a lower signal-to-noise ratio, in the sense that it reflects relatively more exogenous conditions and relatively less the incumbent's skill. This redirection of screening is unambiguously welfare-reducing. Intuitively, screening attains the constrained optimum in the absence of voter bias ($\beta = 0$), and any increase in bias induces a further distortion.

For instance, we might plausibly believe that the noise surrounding economic events is significantly greater than the noise surrounding Presidential actions in the foreign-policy arena. From this perspective, Proposition 2 implies that biased voters overweight skills and outcomes in the noisy economic arena relative to foreign policy related skills and outcomes. They should be electing the chief US diplomat, but the FAE leads them to seek instead a skilled manager of the American economy.

Career concerns determine politicians' incentives as a by-product of their screening. Accordingly, the FAE redirects politicians' incentives exactly as it redirects voters' attention. This redirection of incentives has two separate effects. First, it determines the allocation of productive expenditure across different public services. Second, it determines the total amount of productive expenditure and the level of rent extraction.

To distinguish these two components, we can define the share of public service g over total productive expenditure:

$$\xi_g \equiv \frac{x_g}{b-r}. \tag{9}$$

This equilibrium share must be compared to the optimal share implied by Equation 5. The optimal budget allocation reflects voters' valuation of different public services and returns to spending in their provision. However, it does not reflect differences in volatility, and in particular it is independent of the relative importance of skill and exogenous conditions. As a consequence, rational voters' tendency to skew politicians' incentives towards providing public services with a high signal-to-noise ratio is welfare reducing. Intuitively, rational voters reward politicians who strut their stuff rather than tending to useful but unglamorous issues. The FAE reduces this distortion and as a result it unambiguously makes the allocation of expenditure across public services more efficient.

Proposition 3 *The FAE increases the equilibrium share of productive expenditure devoted to a public service if and only if its provision depends relatively more on exogenous conditions and relatively less on government ability than the provision of other public services:*

$$\frac{\partial \xi_g}{\partial \beta} > 0 \textit{ if and only if } \sum_{j=1}^{G} (\nu_g - \nu_j)\, \alpha_j \rho_j > 0.$$

This budget reallocation increases voters' welfare ($\sum_{g=1}^{G} \alpha_g \rho_g \partial \ln \xi_g / \partial \beta \geq 0$).

The FAE blunts voters perception of differences in noisiness across public services. In reality, the provision of some public services reflects politicians' skills very tightly, while the provision of others depends mainly on exogenous conditions. Biased voters perceive both issues as more informative than they truly are. However, the mistake is naturally lower for services whose true informativeness is high, so little room remains to overestimate it.

Proposition 2 showed that voters' failure to differentiate policy areas according to their signal-to-noise ratio makes political selection less effective. Proposition 1 shows that, conversely, it makes political incentives more effective. The incentives for politicians to allocate resources where they are a better signal-jamming instrument declines. Instead, they tend to be allocated where they are most needed. In the limit as voters are fully biased, the equilibrium allocation of Proposition 1 reaches the optimum:

$$\lim_{\beta \to 1} \frac{x_{g,1}}{x_{j,1}} = \frac{\alpha_g \rho_g}{\alpha_j \rho_j} = \frac{x^*_{g,1}}{x^*_{j,1}}. \tag{10}$$

Returning to the difference between foreign diplomacy and economic policy, Proposition 1 highlights the silver lining of voters' tendency to overestimate the impact of Presidential decisions on the economy. While it tends to reward the lucky instead of the capable, it prompts all administrations to focus on important domestic issues instead of grandstanding on the international stage.

The impact of the FAE on aggregate rent extraction is ambiguous because it reflects two mechanisms. On the one hand, if voter attention is redirected towards public services with higher returns to spending, then politicians refrain from rent extraction because the same amount of spending has a stronger impact on electoral success. On the other hand, if voter attention is redirected towards public services whose provision is more volatile, then politicians indulge in rent extraction because their ability to control their own electoral success is diminished.

To capture formally the balance between the two forces, define the spending shifter: $\psi_g \equiv \nu_g / [1-(1-\beta)\nu_g]$ such that $\partial \ln \xi_g / \partial \beta - \partial \ln \xi_j / \partial \beta = \psi_g - \psi_j$ for any pair of public services g and j. Define the electoral riskiness $\varsigma_g^2 \equiv \{\alpha_g [1-(1-\beta)\nu_g] \sigma_g\}^2$, which equals the variance of the contribution of service g to the incumbent's re-election. Letting $\overline{\mathbb{E}}$ denote the sample mean across public services and $\overline{\mathrm{Cov}}$ the sample covariance, we can establish the following result.

Proposition 4 *The FAE increases rent extraction if and only if it redirects spending towards public services characterized by higher electoral riskiness and lower equilibrium expenditure:*

$$\frac{\partial r}{\partial \beta} \geq 0 \Leftrightarrow \frac{\overline{\mathrm{Cov}}\left(\psi_g, \varsigma_g^2\right)}{\overline{\mathbb{E}}\left(\varsigma_g^2\right)} \geq \frac{\overline{\mathrm{Cov}}\left(\psi_g, \xi_g\right)}{\overline{\mathbb{E}}\left(\xi_g\right)}.$$

The FAE reduces rent extraction if the only difference across public services is in their signal-to-noise ratio ($\partial r / \partial \beta \leq 0$ if $\alpha_g = \alpha$, $\rho_g = \rho$ and $\sigma_g = \sigma$ for all g).

Rent extraction tends to fall if the FAE redirects government spending to public services that attract a greater share of productive expenditure ($\overline{\mathrm{Cov}}\left(\psi_g, \xi_g\right) > 0$). These are services for which signal-jamming through productive investment is particularly appealing for the politician because it has a high expected return. A marginal increase in expenditure on these services translates into a large increase in inferred ability (high ρ_g). Thus, a shift in voter attention towards these services sharpens career concerns.

On the other hand, rent extraction tends to rise if the FAE redirects government spending to public services whose provision is very volatile ($\overline{\mathrm{Cov}}\left(\psi_g, \varsigma_g^2\right) > 0$). These are services for which signal-jamming through productive investment is particularly unappealing for the politician because it has high riskiness. Small changes in expenditure are likely to be dwarfed by large swings in realized ability and exogenous circumstances (high σ_g^2). Thus, a shift towards these services blunts politicians' career concerns, just as an increase in the variance of noise or ability does.

It is easier to convey the intuition behind Proposition 4 by focusing on special cases in which public services are homogeneous along some dimensions.

Corollary 2 *Suppose all public services are equally important for voters' welfare ($\alpha_g = \alpha$ for all g). Then the FAE increases rent extraction if and only if it redirects voter attention towards public services characterized by higher perceived variance of politicians' ability and lower returns to government spending:*

$$\frac{\partial r}{\partial \beta} \geq 0 \Leftrightarrow \frac{\overline{\mathrm{Cov}}\left(\nu_g, \tilde{\sigma}_g^2\right)}{\overline{\mathbb{E}}\left(\tilde{\sigma}_g^2\right)} \geq \frac{\overline{\mathrm{Cov}}\left(\nu_g, \rho_g\right)}{\overline{\mathbb{E}}\left(\rho_g\right)}.$$

Suppose furthermore that the provision of all public services is equally volatile ($\sigma_g = \sigma$ for all g). Then the FAE decreases rent extraction whenever public services with higher returns *to government spending are more influenced by exogenous conditions and less by politicians' ability:*

$$\frac{\partial r}{\partial \beta} \leq 0 \Leftrightarrow \frac{(1-\beta)\,\overline{\mathrm{Var}}\,(\nu_g)}{1-(1-\beta)\,\overline{\mathbb{E}}\,(\nu_g)} + \frac{\overline{\mathrm{Cov}}\,(\nu_g, \rho_g)}{\overline{\mathbb{E}}\,(\rho_g)} \geq 0.$$

If the only difference across public services is in the relative importance of exogenous conditions and politicians' ability, then the FAE reduces rent extraction ($\partial r/\partial \beta \leq 0$ if $\alpha_g = \alpha$, $\rho_g = \rho$ and $\sigma_g = \sigma$ for all g).

The corollary highlights directly the role of changes in voter attention, which trigger equilibrium changes in government spending. Biased voters pay too much attention to public services that provide noisy signals of ability (high ν_g) and too little attention to public services whose provision is instead highly informative. Incentives to refrain from rent extraction then improve if public services characterized by low informativeness also display high returns to government spending and low perceived variance of politicians' ability.

As the corollary highlights, the perceived variance of politicians' ability is directly related to informativeness, unless the FAE is extreme ($\beta = 1$). The more informative public-service provision, the higher the variance of voters' posterior. In the limit, if there is no noise then the posterior has the same volatility as public-service provision. If on the contrary voters receive no information their posterior coincides with their deterministic prior. This negative correlation between noisiness and posterior variance is one channel through which the FAE always tends to reduce rent extraction.

If there is no difference in aggregate volatility across public services, then the FAE reduces rent extraction unless the noisier issues that voter attention turns to have disproportionately lower returns to public spending. In particular, rent extraction certainly falls if returns to spending are identical across public services.

Returning to our example of foreign diplomacy and domestic economic policy, we could reasonably expect voter bias to promote the overall accountability of the federal administration. Economic fluctuations are no less unpredictable than swings in foreign relations, while domestic policy is presumably more responsive to the amount of resources devoted to it—as reflected in its larger share of the federal budget.

An even more extreme example occurs if voters care about two things only: a public service with positive returns to spending and a fixed politician's attribute, such as ideology or personality, that is independent of spending.

Corollary 3 *Suppose voter welfare depends on two public services only ($G = 2$) and that one of them requires no spending ($\rho_2 = 0$). Then the FAE decreases rent extraction if and only if the relative importance of exogenous shocks is higher for the public service that requires spending ($\partial r/\partial \beta \leq 0$ if and only if $\nu_1 \geq \nu_2$).*

In this extreme two-issue scenario, neither differences in welfare weights (α_g) nor differences in aggregate volatility (σ_g) matter. The FAE simply redirects voter attention towards the issue most affected by exogenous shocks, and away from the issue that truly depends the most on politicians' ability. Rent extraction then declines if exogenous conditions matters most for public services that also depend on government spending.

3.4 Demand for Transparency

Propositions 2, 3 and 4 describe changes in politicians' selection and incentives. The FAE worsens screening, but it may still increase voter welfare by inducing an improvement in the budget allocation and a reduction in rents. However, the impact of the FAE on welfare is unambiguously negative when it is applied to constitutional design. A clearheaded institutional designer, who anticipates the voters' flaws and optimizes their welfare accordingly, must be better than an institutional designer with psychological flaws of any form. In this section, we investigate institutional design around transparency, which can take the form of public accounting requirements, public but independent institutions that vet policies, such as the Congressional Budget Office, and the protection of private institutions that provide politically relevant information, such as the free press.

Voters who suffer from the FAE systematically underestimate the extent of noise in outcomes related to the provision of public services. Consequently, they overestimate their own ability to discern the true impact of government policy without any external intellectual aid. As a result, they misunderstand the value of expert analysis of exogenous economic conditions.

For simplicity, assume that at some cost an assessment mechanism can be created that will perfectly reveal exogenous conditions $\varepsilon_{g,t}$. When such information is available, politicians' ability is perfectly revealed in a rational expectations equilibrium. Voter inference obeys:

$$\mathbb{E}\left(\eta_{g,t}^{i_t}|\eta_{g,t-1}, y_{g,t}, \varepsilon_{g,t}\right) = \ln y_{g,t} - \rho_g \ln \bar{x}_g - \eta_{g,t-1}^{i_t} - \varepsilon_{g,t} = \eta_{g,t}^{i_t} + \rho_g \left(\ln x_{g,t} - \ln \bar{x}_g\right). \quad (11)$$

As a result, rational expectations imply that $\mathbb{E}\left(\eta_{g,t}^{i_t}|\eta_{g,t-1}, y_{g,t}, \varepsilon_{g,t}\right) = \eta_{g,t}^{i_t}$ and that screening attains the first best. The incumbent is going to be re-elected if and only if $\sum_{g=1}^{G} \alpha_g \eta_{g,t}^{i_t} \geq 0$.

As we prove in the appendix, there is a unique rational expectations equilibrium when noise is eliminated so long as politicians fully share voters' bias and believe that the variance of their own ability is $\tilde{\sigma}_g^2$. The removal of noise in the assessment of the incumbent's performance induces the first-best allocation of productive public expenditure across public services. The only remaining friction in political agency is then that politicians keep extracting rents.

Incumbents refrain from rent extraction to improve public outcomes and thereby improve the voters' opinions of current leadership. In standard situation, a higher noise to signal ratio in those public outcomes will reduce the politicians' incentives to invest in better outcomes. Transparency reduces rent extraction if the only difference across public services is in the relative importance of exogenous conditions and politicians' ability ($\alpha_g = \alpha$, $\rho_g = \rho$ and $\sigma_g = \sigma$ for all g), or in the special case of Corollary 3 if the relative importance of exogenous shocks is higher for the public service that requires spending ($\nu_1 \geq \nu_2$).

Yet, more transparency can increase rent extraction by shifting the politicians' attention towards activities that generate weaker electoral returns. For example, if the noisier forms of public output also had the lowest returns to public spending, then transparency would shift spending to those low return activities, which could reduce total public spending. Similarly, if the share of noise was higher in activities with higher total variance, then transparency would shift spending towards those activities, and this shift might also increase rent extraction.

The FAE does not distort voters' perception of the benefits that transparency brings, or doesn't bring, by reducing rent extraction. Both biased voters and unbiased voters correctly perceive the change in rent extraction due to extra transparency. Any difference in the demand for transparency between voters affected and unaffected by the FAE occurs through different perceptions about benefits related to selecting the right politicians.

Intuitively, FAE voters can undervalue transparency because they underestimate their own mistakes in screening. They believe noise is limited and that merely observing public service provision enables them to infer quite accurately the incumbent's ability. Thus, they see little need for expert assessment and are willing to devote too few resources to enhance transparency.

Yet this simple intuition is complicated by the non-monotonic role of noise in our model, since we hold the total variance of noise plus ability shocks constant. When there is almost no noise, then there is little demand for transparency since eliminating noise is largely irrelevant. When there is a great deal of noise, the total variance of ability goes to zero, and so screening is unimportant and there is little benefit of transparency. The FAE causes voters to underestimate the amount of noise, which will increase the demand for transparency if the true amount of noise is moderate or small.

If the true amount of noise is quite large, however, fully rational voters correctly believe that screening is near valueless, since the variation in politicians' ability is minimal. FAE voters conversely think that screening on politicians' ability is important, and so they are actually willing to pay more for transparency. We believe that cases where leaders' ability actually does matter are more important and common in the real world, but mathematically, it is certainly possible that overestimating the variance of politicians' ability can increase the demand for transparency if the level of true noise is sufficiently high.

Formally, let $\mathbb{E}\eta_{g,t-1}^{i_t^*}$ denote the expected ability of politicians that are optimally screened under full transparency and $\mathbb{E}\eta_{g,t-1}^{i_t}$ that of politicians who are suboptimally screened by biased voters in the presence of noise. The true improvement in the welfare value of screening brought about by transparency is then $\sum_{g=1}^{G}\alpha_g\left(\mathbb{E}\eta_{g,t-1}^{i_t^*}-\mathbb{E}\eta_{g,t-1}^{i_t}\right)\geq 0$. However, the FAE biases voters' expectations of the ability of the politicians they re-elect, respectively to $\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t^*}$ and $\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t}$, yielding a biased assessment of the improvement in the welfare value of screening: $\sum_{g=1}^{G}\alpha_g\left(\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t^*}-\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t}\right)\geq 0$. The difference between the true welfare gain from transparency and the voters' biased perception of this gain is then simply the difference $\Delta\equiv\sum_{g=1}^{G}\alpha_g\left(\mathbb{E}\eta_{g,t-1}^{i_t^*}-\mathbb{E}\eta_{g,t-1}^{i_t}-\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t^*}+\tilde{\mathbb{E}}\eta_{g,t-1}^{i_t}\right)$.

We can then characterize distorted demand for transparency as a function of the extent of bias (β) and two composite parameters that summarize the extent and distribution of noise: an average $\bar{\nu}\equiv\sum_{g=1}^{G}\alpha_g^2\nu_g\sigma_g^2/\sum_{g=1}^{G}\alpha_g^2\sigma_g^2\in[0,1]$, appropriately weighted for the welfare value and absolute volatility of each public service; and the identically weighted variance parameter $\zeta\equiv\left(\sum_{g=1}^{G}\alpha_g^2\nu_g^2\sigma_g^2/\sum_{g=1}^{G}\alpha_g^2\sigma_g^2-\bar{\nu}^2\right)\left[\bar{\nu}\left(1-\bar{\nu}\right)\right]^{-1}\in[0,1]$, normalized to equal 1 when variance attains its potential maximum given mean $\bar{\nu}$ and bounds $\nu_g\in[0,1]$.

Proposition 5 *The FAE reduces voters' demand for transparency if and only if the extent of voter bias is large enough ($\Delta>0$ if and only if $\beta>B\left(\bar{\nu},\zeta\right)$) relative to the average amount of noise and its variation across public services ($\partial B/\partial\bar{\nu}>0$ and $\partial B/\partial\zeta>0$). If the average amount of noise and its variation across public services are low enough, any amount* *of voter bias reduces demand for transparency:*

$$\Delta > 0 \text{ } if \text{ } \frac{(1-\bar{\nu})(1-\zeta)}{\sqrt{1-\bar{\nu}+\bar{\nu}\zeta}} > \frac{1}{2}.$$

If both the average amount of noise and its variation across public services are limited, there is no public service with very high noise. Then the true welfare value of reducing noise is high, and the simple intuition holds. Voters underestimate noise, and thus certainly underestimate the value of transparency.

Conversely, if noise is so high that ability is truly irrelevant ($\bar{\nu} = 1$), or its variation so stark that each public service reflects only ability or only noise ($\zeta = 1 \Rightarrow \nu_g \in \{0, 1\}$), its welfare cost would be nil in the absence of bias. The FAE then has two effects. First, it induces voters to believe that noise is costly when it actually isn't. They refuse to believe that ability is invariant and politicians differ only in luck, so they engage in an inference problem that seems hard and is truly impossible. As a result, they tend to overestimate the value of transparency. However, their mistaken belief is also self-fulfilling. Since they believe that noise is costly, they actually make it costly by distorting their screening. They are unaware of this cost of their mistake, and thus they tend to underestimate the value of transparency. As the extent of voter bias increases, so does its cost. Hence it becomes more and more likely that demand for transparency is insufficient on net.

Proposition 5 confirms that this intuition extends smoothly from the limit case to interior parameter values. If voters are fully biased, they always underestimate the value of transparency (if $\beta = 1$ then $\Delta \propto \sqrt{1-\bar{\nu}}\left(1-\sqrt{1-\bar{\nu}}\right) > 0$). When bias is only partial, it continues to imply insufficient demand for transparency so long as noise is not too extreme.

The basic intuition of our result is that FAE voters' may not see the need for experts, because they believe that outcomes speak for themselves. We have framed their demand for expertise as a willingness-to-pay for transparency, which can be interpreted as the extent to which they are willing to fight attempts to squelch the free press. This framework also predicts, as long as noise is not too high, that voters who suffer less from the FAE are more likely to care about transparency. If we were willing to accept that better educated people suffer less from the FAE, then our result could explain why more educated people appear to be more concerned about the freedom of the press in opinion polls (Pew Global Institute 2015).

Alternatively, voters might believe that experts come with their own hidden biases. In that case, the FAE would lead them to put little weight on such expertise relative to observed economic outcomes, because they believe that the bias is worse than the noise. This might explain the scorn for expert opinion displayed by many US voters.

4 The FAE in a Signaling Model of Political Agency

In this section, we apply our interpretation of the Fundamental Attribution Error to classic political signaling models developed by Besley (2007) and ???. In this model, politicians can be always honest, always dishonest, or opportunistic. Formally, this will be modeled by heterogeneity in the personal, perhaps psychic, cost of corruption. Politicians, like U.S. Presidents, serve for at most two terms, and in their first term, opportunistic politicians may refrain from corruption to appear as honest politicians. The FAE means that voters overestimate the share of politicians who are either always honest or dishonest, and underestimate the share of politicians who are opportunists that respond to incentives.

4.1 Preferences and Technology

In period t, the ruling politician i_t allocates an exogenous budget $\hat{r} > 0$ to public-good provision and rent extraction $r_t \in [0, \hat{r}]$. This rent extraction can be unbridled corruption, perhaps through overpayment to connected contractors, or directing funds to pet causes of the leader that are not valued by the voters.

Public goods yield voter welfare

$$u_t = \eta_{i_t} + \hat{r} - r_t. \tag{12}$$

The term η_{i_t} reflects the ability of the ruling politician, which is i.i.d. across politicians and uniformly distributed with mean zero and maximum $\hat{\eta}$. For ease of notation, we will let $F(\eta)$ denote the cumulative distribution function of $U[-\hat{\eta}, \hat{\eta}]$.

Being in power at t gives politician i flow utility

$$v_{i,t} = \upsilon + \gamma_{i,t} r_t. \tag{13}$$

The flow utility from not being in power is normalized to zero. The term υ reflects the general benefit ("ego rent") of holding office. The term $\gamma_{i,t}$ reflects the benefit that politicians get from rent extraction. This benefit is both politician- and time-specific.

Politicians differ in their taste for rent extraction, which is each politician's private information. They belong to three distinct types. Some are Honest and have $\gamma_{i,t} = \check{\gamma} < 0$ for all t. Some are Corrupt and have $\gamma_{i,t} = \hat{\gamma}$ for all t. The remainder are Opportunists and draw an independent realization $\gamma_{i,t}$ every period from an identical distribution with mean $\bar{\gamma}$ and cumulative distribution function $H(\gamma)$ on the interval $[0, \hat{\gamma}]$.

We assume that $\hat{\gamma}$ is high enough that Corrupt politicians always engage in maximum rent extraction ($r_t = \hat{r}$ if $i_t \in C$). We also adopt an equilibrium refinement based on Banks and Sobel's (1987) universal divinity that ensures that Honest politicians never engage in rent extraction ($r_t = 0$ if $i_t \in H$). We provide below formal statements of both assumptions: respectively Equation (22) and Lemma 1.

4.2 Electoral Discipline with Term Limits

We consider an infinitely repeated model in which politicians can be re-elected at most once. Both voters and politicians have a discount factor $\delta \in (0, 1)$.

In a politician's second and final term in office, there are no electoral incentives to influence incumbent behavior. Honest politicians extract no rents and have a value of holding office $v_{H,2} = \upsilon$. Both Corrupt politicians and Opportunists extract maximum rent. Their value of holding office is respectively $v_{C,2} = \upsilon + \hat{\gamma}\hat{r}$ and $v_{O,2} = \upsilon + \gamma_{O,2}\hat{r}$.

A first-term politician is re-elected to a second term according to the following timeline.

1. A randomly drawn first-term politician i_1 allocates the budget to public goods and rent extraction r_1, knowing his type but not his ability.

2. The politician's ability η_{i_1} is realized and publicly observed. The voters observe rent extraction r_1.

3. An election is held, pitting the incumbent against a random challenger drawn from the same pool.

Let W_1 denote voters' expectation of their own welfare when a random politician starts his first term in office. Given rational expectations of second-term behavior from politicians of each type, voter's expected welfare from re-electing the incumbent i_1 is

$$\mathbb{E}\left(W_2|i_2 = i_1\right) = \eta_{i_1} + \hat{r}\pi\left(r_1\right) + \delta W_1, \tag{14}$$

where $\pi\left(r_1\right)$ denotes the voters' posterior assessment of the probability that the incumbent is Honest.

Voters infer a politician's type from his rent extraction, which is publicly observed because so are ability η_{i_1} and welfare u_1, and there is no uncertainty concerning the state of the world. Expected voter welfare from replacing the incumbent with the challenger is W_1. As a consequence, the incumbent is re-elected if and only if $\mathbb{E}\left(W_2|i_2 = i_1\right) > W_1$, namely if and only if

$$\eta_{i_1} > \left(1-\delta\right)W_1 - \hat{r}\pi\left(r_1\right). \tag{15}$$

For an unknown realization of ability, the probability of re-election conditional on rent extraction then equals

$$p\left(r_1\right) = 1 - F\left(\left(1-\delta\right)W_1 - \hat{r}\pi\left(r_1\right)\right). \tag{16}$$

Trading off current rent extraction and the desire to get re-elected and extract rents in the future, the incumbent chooses

$$r_1 = \arg\max_{r}\left\{\upsilon + \gamma_{i,1}r + p\left(r\right)\mathbb{E}v_{i,2}\right\}. \tag{17}$$

4.3 Equilibrium

We consider perfect Bayesian equilibria in which voters apply Bayes rule correctly, although with biased priors. Furthermore, we select equilibria that satisfy the following refinement based on Banks and Sobel's (1987) universal divinity.

Assumption 1 *For any off-equilibrium level of rent extraction r, define $\bar{p}_i\left(r\right)$ as the re-election probability that makes politician i indifferent between his equilibrium payoff and the payoff from r. Voters ascribe the deviation with certainty to the type of politician with the lowest value of $\bar{p}_i\left(r\right)$.*

This equilibrium refinement suffices to pin down the behavior of Honest politicians.

Lemma 1 *Under Assumption 1, Honest politicians never extract any rents in equilibrium.*

Since Honest politicians dislike extracting rents, in their first term they extract the minimum level of rents observed on the equilibrium path. Our refinement concept ensures that this level must be their preferred one, namely no rent extraction. Otherwise, by deviating to it they could enjoy the reduction in rents while simultaneously signalling their honesty to voters and thus maximizing their probability of re-election.

Given that Honest politicians do not extract rents, even voters with less than full Bayesian rationality understand that any rent extraction in the first period signals that the incumbent politician is not Honest and will maximize rent extraction if re-elected ($\pi(r_1) = 0$ for all $r_1 > 0$). As a consequence, equilibrium rent extraction in the first period can only take an extreme value ($r_1 \in \{0, \hat{r}\}$). Either the incumbent extract no rents in order to seem Honest, or he admits he is not Honest and maximizes rent extraction.

Corrupt politicians choose the latter strategy if they have a sufficiently high desire for rents that they maximize rent extraction whenever they hold office. We assume that $\hat{\gamma}$ is large enough for this to be the case: Equation (22) below provides the formal lower bound.

Denote by $\pi_1 \equiv \pi(0)$ voters' posterior when they observe no rent extraction, and by

$$\eta^* \equiv (1-\delta) W_1 \tag{18}$$

the minimum ability that gets a rent-extracting politician re-elected, which coincides with voters' assessment of their own average flow utility.

Opportunists who anticipate the equilibrium ability thresholds for re-election, respectively η^* for maximum rent extraction and $\eta^* - \pi_1 \hat{r}$ for no rent extraction, prefer to refrain from rent extraction in the first term if and only if their current desire for rents is sufficiently low:

$$\gamma_{i,1} \leq \gamma^* \equiv \delta \left(\frac{\upsilon}{\hat{r}} + \bar{\gamma} \right) \left[F(\eta^*) - F(\eta^* - \pi_1 \hat{r}) \right]. \tag{19}$$

Let voters believe that the prior share of Honest politicians is π_0 and the prior share of opportunists is κ. These priors need not coincide with the true population shares because of the FAE. If voters expect equilibrium behavior with a threshold γ^*, they infer the posterior probability

$$\pi_1 = \frac{\pi_0}{\pi_0 + \kappa H(\gamma^*)}. \tag{20}$$

As a result, their expectation of their own welfare when electing a random politician to his first term implies that η^* satisfies the recursive definition:

$$\begin{aligned} \eta^* = {} & \left[1 - \pi_0 - \kappa H(\gamma^*)\right] \delta \int_{\eta^*}^{\infty} (\eta - \eta^*) \, dF(\eta) \\ & + \left[\pi_0 + \kappa H(\gamma^*)\right] \left[\hat{r} + \delta \int_{\eta^* - \pi_1 \hat{r}}^{\infty} (\eta - \eta^* + \pi_1 \hat{r}) \, dF(\eta) \right], \end{aligned} \tag{21}$$

such that $\eta^* \in (0, \hat{\eta} + \pi_1 \hat{r})$.[3]

A perfect Bayesian equilibrium subject to the FAE and to the refinement in Lemma 1 is jointly defined by Equations (19), (20) and (21). We begin by providing a sufficient condition for equilibrium uniqueness.

[3] The proof of Lemma 3 shows explicitly the steps required to derive Equation (21)

Lemma 2 *If the importance of ability relative to honesty is sufficiently high that*

$$\frac{\hat{\eta}}{\hat{r}} \geq \frac{\pi_0}{\pi_0 + \kappa H\left(\frac{1}{2}\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\right)},$$

there is a unique equilibrium. In equilibrium, Opportunists sometimes refrain from rent extraction in the first period ($\gamma^* > 0$) and rent-extracting politicians are sometimes dismissed even if their ability is above average ($\eta^* > 0$).

Could there be one equilibrium in which few opportunists act honestly and a second equilibrium in which many opportunists act honestly?

In an equilibrium where more opportunists act honestly in the first period, voters come to believe that acting honestly does not mean being honest. This inference directly reduces the probability of re-election for politicians who act honestly. Moreover, an equilibrium in which opportunists act honestly in their first term increases the returns to electing a new leader. The higher expected value of a first-term politician adds an indirect effect that further reduces the probability of being re-elected conditional upon behaving honestly. Formally, the sum of these two effects is reflected in Equation (21), which defines voters' screening of politicians (η^*) as an increasing function of Opportunists' pooling (γ^*).

The potential for multiplicity emerges because a greater share of Opportunists acting honestly in their first term also reduces the probability of re-electing politicians who behave dishonestly. If the greater appeal of a first term politician reduces the probability of re-electing a dishonest politician more than it reduces the probability of re-electing a politician who behaved honestly, then there is the kind of strategic complementarity that generates multiplicity. An extreme example of this effect occurs when all politicians who are honest are re-elected in either equilibrium. Politicians who behave dishonestly are re-elected at a high rate in the equilibrium with few opportunists behaving honestly in their first term, and at a low rate in the equilibrium with many opportunists behaving honestly.

Our assumption about the importance of ability relative to honesty rules out such multiple equilibria. Formally, Lemma 2 ensures that the least capable politicians are always dismissed even if they behave honestly in the first term ($F(\eta^* - \pi_1\hat{r}) > 0$ for all $\eta^* \geq 0$). Then, Equation (19) describes Opportunists' incentives to pool (γ^*) as a weakly decreasing function of voters' screening (η^*).

Voters' trade-off between ability and honesty is intuitively more likely to favor the former when variance in ability ($\hat{\eta}$) is higher and the potential for rent extraction ($\hat{r}$) is lower; but also when opportunism is more prevalent and bias less extreme (high $\tilde{\kappa}$), when honesty is less prevalent (low π_0), patience (δ) is higher and politicians are keener on holding office regardless of rent extraction (high υ).

Proposition 6 *There is a threshold $\Xi > 0$ such that, if the importance of ability relative to honesty is sufficiently high that $\hat{\eta}/\hat{r} \geq \Xi$, then there is a unique equilibrium, in which the extent of pooling is defined by γ^* such that*

$$\gamma^* = \frac{1}{2}\delta\frac{\upsilon + \bar{\gamma}\hat{r}}{\hat{\eta}}\frac{\pi_0}{\pi_0 + \kappa H(\gamma^*)},$$

while the extent of screening is defined by $\eta^* \in (0, \hat{\eta}]$ *such that*

$$\eta^* = [1 - \pi_0 - \kappa H(\gamma^*)] \frac{\delta}{4\hat{\eta}} (\hat{\eta} - \eta^*)^2 + [\pi_0 + \kappa H(\gamma^*)] \left\{ \hat{r} + \frac{\delta}{4\hat{\eta}} \left[\hat{\eta} - \eta^* + \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)} \right]^2 \right\}.$$

Intuitively, if ability is sufficiently important relative to honesty there is positive probability that a politician is capable enough to be re-elected even if he extracted rents ($\eta^* < \hat{\eta}$). The amount of equilibrium screening is then interior in every sense, since the condition in Lemma 2 also ensures that there is positive probability that a politician is incapable enough to be dismissed even if he did not extract rents ($\eta^* - \pi_1 \hat{r} > -\hat{\eta}$). Within its support, a uniform distribution of ability then ensures that (19) determines the equilibrium amount of pooling independently of the equilibrium amount of screening.

The condition that ensures that Corrupt politicians extract maximum rents is simply $\hat{\gamma} \geq \gamma^*$, namely:

$$\hat{\gamma} \geq \frac{1}{2} \delta \frac{\upsilon + \bar{\gamma} \hat{r}}{\hat{\eta}} \frac{\pi_0}{\pi_0 + \kappa}. \tag{22}$$

The probability that voters re-elect a politician who refrained from extracting rents in his first term is

$$p_0 \equiv \frac{1}{2} \left(1 - \frac{\eta^*}{\hat{\eta}} + \pi_1 \frac{\hat{r}}{\hat{\eta}} \right); \tag{23}$$

and the probability that voters re-elect a politician who extracted rents in his first term is

$$p_r \equiv \frac{1}{2} \left(1 - \frac{\eta^*}{\hat{\eta}} \right). \tag{24}$$

Exploiting Equations (20), (23) and (24), the interior equilibrium of Proposition 6 can be described most compactly by

$$\gamma^* = \frac{1}{2} \delta \frac{\upsilon + \bar{\gamma} \hat{r}}{\hat{\eta}} \pi_1 \tag{25}$$

and

$$\frac{\eta^*}{\hat{\eta}} = \left(1 - \frac{\pi_0}{\pi_1} \right) \delta p_r^2 + \frac{\pi_0}{\pi_1} \left(\frac{\hat{r}}{\hat{\eta}} + \delta p_0^2 \right), \tag{26}$$

noting that by Bayes' rule the probability that a politician refrains from rent extraction is π_0 / π_1, while by the uniform distribution of ability the option value of re-electing a politician is p_r^2 if he extracted rents and p_0^2 if he refrained from rent extraction.

4.4 Comparative Statics

In this section, we derive comparative statics for equilibrium incentives and screening. We are particularly interested in the impact of the FAE, which changes voters' perception of politicians' honesty and opportunism. To capture this formally, we denote the extent of voter bias by β such that priors are $\pi_0(\beta)$ and $\kappa(\beta)$. The FAE reduces perception of opportunism: $\partial \kappa / \partial \beta < 0$. We would expect it to weakly increase the perceived prevalence of both pure types, so that $\partial \pi_0 / \partial \beta \geq 0$. However, our formal analysis can also handle the case of bias that reduces perceptions of both opportunism and honesty, raising only perceptions of corruption.

Before discussing our formal results, we illustrate the intuition of our comparative statics with graphed simulations of the model. The simulations shown in the figures assume no discounting ($\delta = 1$), a uniform distribution of greed on the unit interval ($\gamma_{i,t} \sim U[0,1]$), a unit ratio of maximum ability to maximum rent extraction ($\hat{r} = \hat{\eta}$) and a pure value of holding office $\upsilon = \hat{r}/2$.

Figure 1 illustrates the basic structure of the equilibrium, assuming that one-half of non-Opportunists are thought to be honest ($\pi_0 = (1-\kappa)/2$). The figure shows the changing probability of re-election as a function of the perceived probability that a new leader will be an Opportunist. This probability both determines the returns from seeking a new leader, and the interpretation that voters put on the action of the incumbent. When voters believe that Opportunists are common, then the probability of being re-elected if the incumbent refrains from rent-extraction declines, as shown in the declining red line. This decline reflects, partially, the tendency of voters to attribute such restraint to opportunism, not innate honesty.

The blue line shows the probability that an Opportunist will behave honestly in his first term of honest, which also declines with the perceived share of Opportunists. If the Opportunist cannot credibly convince voters that he is Honest by restraining from corruption, then there is no reason for him to behave honestly. When voters mistakenly believe that no one is Honest, then the incentive to deceive is at its strongest. When voters are cynical and think that good behavior is just opportunism in disguise, then there is little reason to disguise. Notably, voters are actually suffering harm here because of this more realistic outlook, which reduces the tendency of politicians to behave well.

The upward sloping green line shows the probability that politicians who behave badly will get re-elected. This probability converges to the probability that politicians who behave well will get elected as the perceived share of opportunists rises to one. When voters believe that all politicians are intrinsically identical, then behavior makes no difference to re-election.

Figure 2 helps to understand the impact that the fundamental attribution error has on re-election, and this depends on what an increase in κ does to the value of π_0. One possibility is that an increase in the perceived share of Opportunists means that there are fewer true rogues, but that the share of Honest politicians remains unchanged, and in that case $\partial\pi_0/\partial\kappa = 0$. Another possibility is that an increase in the perceived share of Opportunists comes out of the perceived share of Honest politicians and in that case $\partial\pi_0/\partial\kappa = -1$. Given that we assume that the share of non-Opportunists that are Honest equals $1/2$, perhaps a natural middle ground is that $\partial\pi_0/\partial\kappa = -1/2$.

The impact of the FAE on π_0 matters because that determines how the FAE changes beliefs about new leaders. The FAE does not change beliefs about incumbents who have expropriated, since they are all Opportunists or bad types, and will expropriate in the next period. Consequently, the FAE has only impacts the re-election of politicians who have expropriated by changing beliefs about their replacement. The largest impact of the FAE, in this simulation, works through changing the belief that a new leader will expropriate in his first term in office.

When the share of Opportunists is quite low, then Opportunists don't expropriate in the first period in our example. Consequently, as long as $|\partial\pi_0/\partial\kappa|$ is less than one half, so that the FAE doesn't disproportionately generate beliefs about more Honest politicians, then it doesn't damage the re-election prospects for the incumbents. When $|\partial\pi_0/\partial\kappa|$ is close to one, then the FAE generates far more optimism that a newly elected leader will be Honest, and this increase the appeal of electing a new leader.

When the share of Opportunists increases, Opportunists expropriate more often in the first period, and this increases the voters' desire to have the Honest type relative to Opportunists. Consequently, the FAE will hurt incumbents who have cheated for a wider range of $|\partial\pi_0/\partial\kappa|$, since even a small increase in the number of Honest politicians makes it more appealing to take a risk with a new leader. The FAE is more likely to yield an incumbency advantage for dishonest politicians when people start with a belief that the share of Opportunists is low.

With incumbent politicians who have behaved honestly, the FAE also leads voters to think that the incumbent is truly honest, rather than an opportunist. This effect always increases the FAE's impact on the incumbency advantage, which is shown by the red line always being above the blue line. Mathematically, this effect increases the direct effect of beliefs about κ on incumbency and mutes the indirect effect of $\partial\pi_0/\partial\kappa$.

As the perceived share of leaders who are Opportunists rises, the direct effect of beliefs about κ on incumbency and the indirect effect of $\partial\pi_0/\partial\kappa$ both decline, partially because Opportunists pool less often, and is proportional to $(\kappa\partial\pi_0/\partial\kappa - \pi_0)\, H(\gamma^*)$. Initially, the direct effect of beliefs about κ declines more swiftly than the indirect effect of $\partial\pi_0/\partial\kappa$ and consequently the red line falls. Yet the impact of $\partial\pi_0/\partial\kappa$ reaches zero at a value of $\kappa < 1$, while the direct effect of κ does not, and at that point, the red line starts increasing dramatically eventually going to infinity. After that point, the FAE can only increase the incumbency advantage, but the effect is small.

Figures 3 and 4 repeat Figures 1 and 2, but increase the share of non-Opportunists who are thought to be Honest to eighty percent. This shift sharply reduces re-election probabilities for incumbents when the share of Opportunists is thought to be low. A higher probability of Honest replacements naturally reduces the appeal of incumbents. When the share of Opportunists is thought to be high, then this effect is muted since the honesty of non-Opportunists is less important when non-Opportunists are rare. The pooling probabilities are also higher, since the importance of appearing to be Honest is even larger, and because voters are more likely to think that a politician who looks Honest is Honest. In this case, the impact of the perceived share of Opportunists on the probability of re-election for politicians who behave honestly is non-monotonic. Increases in the share who are Opportunist strongly reduces the appeal of new leaders in this case.

Figure 4 shows that this shift makes it even less likely that the FAE will help incumbents who have behaved dishonestly. These politicians are very unlikely to be re-elected, since the non-incumbents are so likely to be Honest. The FAE really hurts them because one impact of the FAE is to increase the value of new leader who behaves well and such leaders are particularly common when the share of honest leaders is so high. The power of the FAE on the incumbency advantage of leaders who behave well is slightly muted in this scenario, but the shape of the curve is quite similar to Figure 2.

Figures 5 and 6 show results when only 20 percent of non-Opportunists are thought to be Honest. Naturally, this raises the re-election possibility of both types of leaders considerably. In this case, the FAE is extremely likely to improve the incumbency advantage for leaders who have behaved honestly. Since dishonest types don't restrain, the FAE always increases the belief that pooling types are Honest, and this is extremely value when non-incumbents are largely thought to be dishonest.

Figures 7 and 8 examine welfare as a function of bias. In this case, we assume that individuals believe that the share of Opportunists equals $1 - \beta$ times the true share of Opportunists. The level of bias does not impact their beliefs about the honesty of non-Opportunists. In Figure 7, we assume that in reality one-half of the leaders are Opportunists. In Figure 8, we assume that ninety percent of leaders are really Opportunists. In both cases, one-half of non-Opportunists are thought to be Honest.

In both cases, the amount of pooling increases with the amount of bias. The impact is naturally much larger when most politicians really are Opportunists. This increase in pooling reflects the fact that bias increases the gap in re-election probabilities between the politicians who refrain from extraction and the politicians who don't refrain. Bias, after all, increases, the belief that a politician who refrains is actually Honest.

Welfare is increasing with bias. This finding is not universal. It is possible that erroneous beliefs about incumbents end up causing more harm than good, but for most reasonable parameter values, in this setting, the FAE helps voters out. Since the FAE increases the incentives to behave well, politicians are less likely to behave badly and voters benefit. This effect is particularly strong when most politicians are Opportunists. A correct skepticism just has the impact of ensuring that politicians will behave badly since well-informed voters will know that they are all Opportunists anyway.

Figure 9 and 10 explore heterogeneity in beliefs about the honesty of non-opportunists, under the assumption that 90 percent of leaders are actually Opportunists. The first notable difference between the graphs is that welfare is significantly higher, when there is no bias, when the majority of non-Opportunists are actually Honest. This welfare gap reflects both the direct benefit of added honesty, and the much larger benefit of inducing significant pooling on the part among the Opportunists.

The welfare benefits of bias are actually higher in the case where more non-Opportunists are dishonest. Pooling is always lower in this case, but the impact of bias on pooling is larger. Dishonest politicians never refrain, so the large share of dishonest politicians doesn't reduce the benefit to Opportunists from pooling. In both cases, pooling goes to 100 percent as bias goes to 100 percent.

Such error is not always benign. Figure 11 illustrates cases in which the FAE can be harmful. For this figure we return to the assumptions of Figure 8 (90% Opportunists and 10% Honest politicians) but we increase the advantages of just being in office, relative to expropriation ($\upsilon = 3\hat{r}/2$ instead of $\upsilon = \hat{r}/2$). Those benefits have the effect of inducing pooling among Opportunists at lower levels of bias. Once pooling is complete, then the FAE's impact turns negative, since it has no more positive incentive effects and it works primarily to induce worse re-election decisions.

In the interior equilibrium described by Proposition 6, Opportunists' incentives to refrain from rent extraction in their first term have fully unambiguous comparative statics.

Proposition 7 *Opportunists are less likely to extract rents in their first term when they are more patient ($\partial\gamma^*/\partial\delta > 0$), holding office is more valuable ($\partial\gamma^*/\partial\upsilon > 0$), the scope for rent extraction is higher ($\partial\gamma^*/\partial\hat{r} > 0$) and politicians' ability is less variable ($\partial\gamma^*/\partial\hat{\eta} < 0$).*

Opportunists' incentives to refrain from rent extraction improve with voters' prior for honesty ($\partial\gamma^*/\partial\pi_0 > 0$) and worsen with voters' prior for opportunism ($\partial\gamma^*/\partial\kappa < 0$).

Proposition 7 shows that the basic comparative statics in most political signaling models are unchanged by the FAE. Opportunists refrain from rent extraction to increase their probability of re-election, so their incentives improve with their intrinsic motivation to hold office and with their patience, which makes them keener on delayed gratification. When the scope for rent extraction is higher and politicians' ability less variable, voters place more weight on rent extraction and less on ability: as a result, Opportunists' incentives become sharper.

Proposition 8 *The FAE improves Opportunists' incentives to refrain from rent extraction unless it causes a greater proportional fall in perceived honesty than in perceived opportunism ($\partial\gamma^*/\partial\beta > 0$ if and only if $-\partial\ln\kappa/\partial\beta > -\partial\ln\pi_0/\partial\beta$).*

The effect of voters' priors is a direct consequence of Bayesian inference. Incentives are sharper if voters believe there are more Honest politicians and fewer Opportunists, because then pooling with the former gives the latter a higher probability of re-election. The FAE should cause the perceived share of honest politicians to (weakly) rise and the perceived share of opportunists to decline. Both of these effects will cause incentives for honesty to strengthen. The most important implication of Proposition 8 is thus that the FAE always improves Opportunists' incentives because it strictly lowers voters' perception of opportunism while weakly raising their perception of honesty.

The comparative statics for the screening of rent-extracting politicians have only modest ambiguity, concerning changes in voters' priors.

Proposition 9 *Politicians who extract rents in their first term are less likely to be re-elected when people are more patient ($\partial p_r/\partial\delta < 0$), holding office is more valuable ($\partial p_r/\partial\upsilon < 0$), the scope for rent extraction is higher ($\partial p_r/\partial\hat{r} < 0$) and politicians' ability is less variable ($\partial p_r/\partial\hat{\eta} > 0$). The probability of re-election of politicians who extracted rents declines with voters' prior for honesty, and more weakly with voters' prior for opportunism ($\partial p_r/\partial\pi_0 < \partial p_r/\partial\kappa < 0$).*

Politicians' pure office-seeking motivation influences screening only through the indirect effect of changes in incentives. When holding office is more valuable, Opportunists are less likely to extract rents in their first term. Voters understand that electoral incentives are more powerful, and as a consequence they are less willing to re-elect politicians to a second term as lame ducks, preferring instead a fresh challenger.

This indirect effect is also present when patience or the relative importance of rent extraction and ability vary. In those cases, however, it is also reinforced by a direct effect. When voters are more patient, they are keener on the option value of electing a challenger who can be re-elected if successful. When the potential for rent extraction is higher and the variance of ability lower, voters are naturally less willing to tolerate rent extraction albeit accompanied by high ability.

Proposition 10 *There is a threshold $\Psi_r > 1$ such that the FAE makes politicians who extract rents in their first term less likely to be re-elected if and only if it induces a decline in perceived opportunism not larger than Ψ_r times the increase in perceived honesty ($\partial p_r/\partial\beta < 0$ if and only if $-\partial\kappa/\partial\beta < \Psi_r\partial\pi_0/\partial\beta$).*

Voters' priors determine re-election probabilities through the sum of two effects. First, voters are keener to dismiss the incumbent when they perceive a greater probability of finding a challenger who won't extract rents in his first term. It rises more weakly with the prior for opportunism, since some but not all Opportunists will refrain from rent extraction. Second, voters are keener to dismiss a rent-extracting incumbent when they perceive a higher option value of finding a challenger who instead won't extract rents. This option value is increasing in the prior for honesty but decreasing in the prior for opportunism, since it reflects the posterior probability that a politician who refrains from rent extraction in his first term is indeed Honest and not merely an Opportunist.

The combination of these two effects implies that screening rises more sharply with the prior for honesty than that for opportunism, but that it rises with the latter as well. Intuitively, since the incumbent has been revealed as a rent extractor, voters prefer the challenger whenever he's less likely to be Corrupt. Even if he's an Opportunist, at least he need not extract rents in his first term.

The FAE reduces voters' prior for opportunism. Proposition 9 shows that its impact on the reelection of rent extractors then depends on how it changes the prior for honesty. In one limit case, if voters only mistake opportunism for corruption in the pool of challengers, then bias always makes them pessimistic and hence more tolerant of rent-extracting incumbents ($\partial\pi_0/\partial\beta = 0 \Rightarrow \partial p_r/\partial\beta > 0$). In the opposite limit case, if voters only mistake opportunism for honesty in the pool of challengers, then bias always makes them optimistic and hence less tolerant of rent-extracting incumbents ($\partial\pi_0/\partial\beta = -\partial\kappa/\partial\beta \Rightarrow \partial p_r/\partial\beta < 0$). In general, optimism about challengers prevails as long as the prior for honesty increases enough as the prior for opportunism declines. Since the effect of the prior for honesty is stronger, its increase can compensate a decrease in the prior for opportunism that is larger by a factor $\Psi_r > 1$.

Greater ambiguity emerges in the comparative statics for the screening of politicians who refrain from rent extraction. We know from Proposition 7 how the difference $p_0 - p_r$ varies, but when its variation has the opposite sign as that in p_r the net effect can be either an increase or a decline in p_0.

Proposition 11 *Politicians who do not extract rents in their first term are less likely to be re-elected when people are more patient ($\partial p_0/\partial\delta < 0$) and holding office is more valuable ($\partial p_0/\partial\upsilon < 0$). The probability of re-election of politicians who did not extract rents declines with voters' prior for opportunism ($\partial p_0/\partial\kappa < 0$).*

As in Proposition 9, comparative statics reflect two channels: changes in the perceived probability that the challenger will refrain from rent extraction, and changes in the option value of re-electing a politician who did not extract rents. The key difference in Proposition 11 is that a higher option value lowers the probability of reelecting a rent extractor, but raises the probability of reelecting a politician who refrained from rent extraction. As a result, the latter probability is sharply declining in perceived opportunism. Not only does the pool of challengers look more attractive, but at the same time the incumbent is viewed with more skepticism since voters consider him less likely to be Honest.

Instead, the importance of rent extraction relative to ability ($\hat{r}/\hat{\eta}$) and perceived honesty (π_0) have countervailing effects. On the one hand, they make the average challenger more appealing. On the other, they also make the incumbent more appealing, given that he refrained from rent extraction and may thus be truly Honest.

Overall, incumbents who did not extract rents in the first term are less likely to be re-elected when rent extraction is more important relative to ability ($\partial p_0/\partial(\hat{r}/\hat{\eta}) < 0$) if and only if

$$\pi_1 < \frac{\pi_0}{\pi_1} - \delta\left(1 - \frac{\pi_0}{\pi_1}\right) p_r \pi_1. \tag{27}$$

The left-hand side is the probability that the incumbent refrains from rent-extraction in his second-term. On the right-hand side, the first term is the probability that the challenger refrains from rent extraction in his first term. The second term reflects the decline in the option value of electing the challenger. He could extract rents (with probability $1 - \pi_0/\pi_1$) and yet gain re-election (with probability p_r). Voters' expected future value would then decline in proportion to the posterior probability (π_1) that a politician who instead refrained from rent extraction in his first term would continue doing so in his second.

Likewise, changes in voters' posteriors have ambiguous effects.

Proposition 12 *There is a threshold $\Psi_0 < \Psi_r$ such that the FAE makes politicians who do not extract rents in their first term more likely to be re-elected if and only if it induces a decline in perceived opportunism larger than Ψ_0 times the increase in perceived honesty ($\partial p_0/\partial\beta > 0$ if and only if $-\partial\kappa/\partial\beta > \Psi_0\partial\pi_0/\partial\beta$).*

An increase in perceived honesty decreases the probability of re-election for politicians who did not extract rents ($\partial p_0/\partial\pi_0 < 0$ such that $\Psi_0 > 0$) if and only if

$$\left[1 + \delta\left(1 - \frac{\pi_0}{\pi_1}\right) p_r\right] \frac{\hat{r}}{\hat{\eta}} (1 - \pi_1) < \left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right] \frac{\pi_0}{\pi_1} \left[1 + \frac{\kappa\gamma^* h(\gamma^*)}{\pi_0}\right]. \tag{28}$$

The left-hand side reflects the impact of increases in the posterior probability of honesty (π_1). As it rises, the incumbent directly becomes more appealing. Moreover, as we just discussed, the option value of electing the challenger declines, because he may turn out to be dishonest yet competent enough to win re-election. The two effects are captured by the first term in square brackets, and they naturally scale with the importance of rent extraction ($\hat{r}/\hat{\eta}$). The last term in parentheses captures the derivative of the posterior with respect to the prior for honesty, through both the direct effect of an increase in honesty (π_0) and the indirect effect of the induced improvement in Opportunists' incentives (γ^*).[4]

The right-hand side captures the countervailing increase in the challenger's appeal due to the rising probability he will refrain from first-term rent-extraction. The first term in brackets is the expected present value of a politician who refrains from rent extraction: the sum of rents avoided in the first term ($\hat{r}/\hat{\eta}$) and the discounted option value of being able to re-elect such a politician rather than a rent extractor ($\delta\left(p_0^2 - p_r^2\right)$). The second term is the probability that a politician refrains from first-term rent extraction (π_0/π_1). The last term in brackets captures its semi-elasticity with respect to perceived honesty, again taking into account the indirect effect through improving incentives.[5]

[4] Formally, the derivative is $\partial\pi_1/\partial\pi_0 = \zeta(1 - \pi_1)$ with a scaling factor $\zeta \equiv 1/\left[\pi_0 + \kappa H(\gamma^*) + \kappa\gamma^* h(\gamma^*)\right]$ that cancels out on both sides of Equation (28).

[5] Formally, the semi-elasticity is $\partial \ln(\pi_0/\pi_1)/\partial\pi_0 = \zeta\left[1 + \kappa\gamma^* h(\gamma^*)/\pi_0\right]$ with the same scaling factor ζ that cancels out on the left-hand side of Equation (28).

If Equation (28) holds, then the analysis of the FAE in Proposition 11 has the same structure as those in Proposition 9. Since the FAE reduces perceived opportunism, it reduces the screening of politicians unless it is accompanied by a sufficiently large increase in perceived honesty. If voters only mistake opportunism for corruption in the pool of challengers, then bias induces such pessimism that it always increases the incumbent's probability of re-election ($\partial\pi_0/\partial\beta = 0 \Rightarrow \partial p_0/\partial\beta > \partial p_r/\partial\beta > 0$).

However, the two propositions jointly imply that screening of rent-extractors can become stricter while screening of politicians who did not extract rents becomes laxer ($\Psi_0 < \Psi_r$ enables $\partial p_0/\partial\beta > 0 > \partial p_r/\partial\beta$). This possibility is intuitive, because we know from Proposition 7 that the FAE increases Opportunists' incentives by raising the difference between the probability of re-election conditional on no rent extraction and the one conditional on positive rent extraction ($\partial p_0/\partial\beta > \partial p_r/\partial\beta$). If Equation (28) does not hold, voter bias necessarily increases the probability that politicians are re-elected when they refrain from extracting rents, no matter how small the induced increase in perceived honesty.

We can resolve the ambiguity in comparative statics for limit cases of voter beliefs.

Corollary 4 *Suppose voters believe no politicians are Opportunists. Then politicians who do not extract rents in their first term are more likely to be re-elected when the scope for rent extraction is higher (*$\lim_{\kappa\to 0} \partial p_0/\partial\hat{r} > 0$*), politicians' ability is less variable (*$\lim_{\kappa\to 0} \partial p_0/\partial\hat{\eta} < 0$*) and voters' prior for honesty is lower (*$\lim_{\kappa\to 0} \partial p_0/\partial\pi_0 < 0$ *such that* $\lim_{\kappa\to 0} \Psi_0 > 0$*).*

When voters do not believe in opportunism, they are certain (perhaps incorrectly) that an incumbent who did not extract rents is truly Honest ($\lim_{\kappa\to 0} \pi_1 = 1$). As rent extraction becomes more important relative to skill, voters become more likely to re-elect him and more reluctant to roll the dice on a fresh challenger, who could be more or less capable, but only more dishonest, not less. On the other hand, if the prior for honesty increases its only effect is to make the challenger more appealing, because the incumbent cannot seem any more Honest than he already does.

For the opposite limit case, we specify a uniform distribution of Opportunists' greed, so Equation (28) simplifies by $\gamma^* h(\gamma^*) = H(\gamma^*)$, whose limit behavior is unambiguous.

Corollary 5 *Suppose the distribution of Opportunists' greed is uniform (*$\gamma_{i,t} \sim U[0,\hat{\gamma}]$*) and that voters believe no politicians are Honest. If and only if*

$$\sqrt{1+\delta} < \frac{1}{2}\frac{\delta}{\hat{\eta}}\left(\frac{\upsilon}{\hat{\gamma}} + \frac{\hat{r}}{2}\right)\kappa,$$

then politicians who do not extract rents in their first term are more likely to be re-elected when the scope for rent extraction is lower ($\lim_{\pi_0\to 0} \partial p_0/\partial\hat{r} < 0$*), politicians' ability is more variable (*$\lim_{\pi_0\to 0} \partial p_0/\partial\hat{\eta} > 0$*) and voters' prior for honesty is lower (*$\lim_{\pi_0\to 0} \partial p_0/\partial\pi_0 < 0$ *such that* $\lim_{\pi_0\to 0} \Psi_0 > 0$*).*

When voters do not believe in Honesty, Proposition 6 implies that Opportunists have no incentives ($\lim_{\pi_0\to 0} \gamma^* = 0$) because voters never infer honesty ($\lim_{\pi_0\to 0} \pi_1 = 0$). Then voters believe that all politicians always extract rents ($\lim_{\pi_0\to 0} (\pi_0/\pi_1) = 0$), so their re-election is independent of rent extraction ($\lim_{\pi_0\to 0} (p_0 - p_r) = 0$). Politicians are still positively screened, not because they might be honest but because re-electing a lame duck sacrifices the option value of retaining an even more capable challenger ($\lim_{\pi_0 \to 0} p_r = \left(\sqrt{1+\delta} - 1\right)/\delta < 1/2$ for all $\delta > 0$).

In the limit, Equation (27) thus simplifies to:

$$\sqrt{1+\delta} \lim_{\pi_0 \to 0} \pi_1 < \lim_{\pi_0 \to 0} \frac{\pi_0}{\pi_1}. \tag{29}$$

This is also the limit case of Equation (28) when greed is uniformly distributed. Even if the incumbent refrained from rent extraction in his first term, there is a vanishing probability that he will refrain from rent extraction in his second term. However, there is also a vanishing probability that the challenger will refrain from rent extraction in his first term. The question is which vanishes faster.

Corollary 5 formalizes the intuition that good behavior tends to vanish faster in the second than in the first term when there are more Opportunists than purely Corrupt politicians (high κ) and when their initial incentives are stronger because they are patient (high δ), they commonly desire holding office for its own sake rather than to extract rent (high υ and low $\hat{\gamma}$), and voters screen more for honesty than for ability (high $\hat{r}$ and low $\hat{\eta}$).

When this is the case, both an increase in the relative importance of rent extraction and a marginal increase in perceived honesty must have a greater impact on perceptions of the challenger—who would serve a first term—than of the incumbent—who would only serve a second term. As a result, they lower the incumbent's probability of re-election. If instead Opportunists' incentives are relatively low-powered, the condition in Corollary 5 fails and increases in perceived honesty or potential rent extraction make incumbents more likely to gain re-election.

4.5 Demand for Dictatorship

Propositions 9 and 11 consider absolute changes in the screening of politicians. No less important is the extent to which the FAE induces mistakes in screening. Although such mistakes occur constantly on the equilibrium path, their welfare cost emerges most starkly in moments of constitutional crisis.

Intuitively, the FAE leads voters to think that the current good performance of an incumbent represents a permanent characteristic of that politician. Consequently, they should be more enthusiastic about turning a temporarily elected leader into a permanent, dictatorial ruler. This logic is clear when considering an extreme case, in which everyone is actually an Opportunist, but voters believe that a share of these politicians are Honest.

If voters saw things clearly, they would know that a dictator will start behaving badly as soon as electoral incentives are eliminated, and consequently they would always oppose dictatorship. However, with the FAE, good performance is seen as a permanent trait and hence, the demand for dictatorship will be robust.

This simple intuition is compromised in our more general model, because the FAE also increases the expected return from electing a new first-term leader. Consequently, results are more ambiguous than this simple extreme logic would suggest. A simple extension of our baseline model allows us to analyze formally voters' demand for dictatorship.

Suppose that at the end of every politician's first term there is probability $\varepsilon \in (0,1)$ of a constitutional crisis. When such a crisis happens, citizens have the opportunity of choosing to appoint the incumbent as a perpetual dictator. If they refuse, the opportunity passes and the incumbent contests a regular election.

So long as the probability of a constitutional crisis is small enough, we can prove that its only impact on our baseline model is an increase in the effective discount factor.

Lemma 3 *Let time preferences be characterized by the pure discount factor $\delta_0 \in (0,1)$. If the probability of a constitutional crisis is sufficiently low that*

$$\varepsilon \leq \left(\frac{1-\delta_0}{\delta_0}\right)^2,$$

then the equilibrium admits the characterization above for an effective discount factor

$$\delta = \delta_0 \left(1 - \varepsilon + \frac{\varepsilon}{1-\delta_0}\right) \in (\delta_0, 1].$$

The resolution of a constitutional crisis simply coincides with that of a regular election. Naturally, if voters wish to dismiss the incumbent, a fortiori they don't want to retain him as a perpetual dictator. But suppose instead that citizens prefer electing the incumbent to a second term without electoral incentives rather than immediately replacing him with a challenger. Then they have an even stronger preference for never replacing him. Formally, the value of re-electing a rent-extracting incumbent to a second term is $\eta_i + \delta W_1$ and the value of appointing him as perpetual dictator is $\eta_i / (1-\delta)$. Thus, citizens prefer dictatorship to re-election if and only if $\eta_i > (1-\delta) W_1 \equiv \eta^*$. Likewise for an incumbent who refrained from rent extraction, with an indifference threshold $\eta^* - \pi_1 \hat{r}$.

As a consequence, politicians' incentives are described by Equation (19) with the effective discount factor shown in Lemma 3. Intuitively, the incumbent knows that, if his ability is higher than the threshold determined by his rent extraction, then in the absence of a crisis (with probability $1-\varepsilon$) he will be re-elected to a single second term, but if a crisis happens (with probability ε) he will get to rule forever.

Voter screening of politicians is described by Equation (21) with the same effective discount factor. Intuitively, voters's valuation of the challenger (on the right-hand side) reflects a heightened option value because with probability ε a constitutional crisis will enable voters to retain the challenger forever after a promising first term.

Incumbent screening, demand for dictatorship and voters' perceptions of their own welfare are all identically summarized by the equilibrium value of η^*. Thus, we can characterize voter mistakes by comparing this value to the assessment of an unbiased observer with correct priors π_0^u and κ^u and correct equilibrium posterior:

$$\pi_1^u = \frac{\pi_0^u}{\pi_0^u + \kappa^u H(\gamma^*)}. \tag{30}$$

The unbiased assessment of average flow utility is $\eta^u = (1-\delta) W_1^u$, implicitly defined by:

$$\eta^u = [1 - \pi_0^u - \kappa^u H(\gamma^*)] \, \delta \int_{\eta^*}^{\infty} (\eta - \eta^u) \, dF(\eta)$$
$$+ [\pi_0^u + \kappa^u H(\gamma^*)] \left[\hat{r} + \delta \int_{\eta^* - \pi_1 \hat{r}}^{\infty} (\eta - \eta^u + \pi_1^u \hat{r}) \, dF(\eta) \right]. \quad (31)$$

Under the hypothesis of Proposition 6, this can be rewritten as:

$$\frac{\eta^u}{\hat{\eta}} = \left(1 - \frac{\pi_0^u}{\pi_1^u}\right) \delta p_r \left(1 - p_r - \frac{\eta^u}{\hat{\eta}}\right) + \frac{\pi_0^u}{\pi_1^u} \left[\frac{\hat{r}}{\hat{\eta}} + \delta p_0 \left(1 - p_0 - \frac{\eta^u}{\hat{\eta}} + \pi_1^u \frac{\hat{r}}{\hat{\eta}}\right) \right]. \quad (32)$$

We can then compare biased voters' demand for dictatorship with that of an unbiased welfare maximizer, which is characterized by thresholds η^u and $\eta^u - \pi_1^u \hat{r}$.

Proposition 13 *Suppose biased voters do not underestimate the prevalence of honesty ($\pi_0 \geq \pi_0^u$). Then they have insufficient demand for dictatorship by competent politicians who extracted rents if and only if they sufficiently overestimate politicians' honesty and, less importantly, their opportunism: there is a threshold $\Xi_r > 1$ such that $\eta^* > \eta^u$ if and only if $\Xi_r (\pi_0 - \pi_0^u) + \kappa - \kappa^u > 0$. Biased voters may nonetheless have excessive demand for dictatorship by incompetent politicians who refrain from rent extraction: there is a threshold $\Xi_0 < \Xi_r$ such that $\eta^* - \pi_1 \hat{r} > \eta^u - \pi_1^u \hat{r}$ if and only if $\Xi_0 (\pi_0 - \pi_0^u) + \kappa - \kappa^u > 0$.*

Just as for comparative statics in Proposition 9, the comparison between true and perceived welfare reflects two effects. First, voters are too optimistic about the pool of challengers if they overestimate the prevalence of either honesty or opportunism. Second, voters are too optimistic about the option value of screening an incumbent after his first term if they overestimate honesty or underestimate opportunism. For perceptions of honesty, the two forces are mutually reinforcing. For perceptions of opportunism they counteract each other. Proposition 13 proves, however, that the first effect dominates if (but not only if) voters do not underestimate the prevalence of honesty.

Thus, the FAE tends to make voters too pessimistic through a decline in perceived opportunism ($\kappa < \kappa^u$). Overall, it may nonetheless end up making them too optimistic if perceived honesty rises enough to compensate. In one limit case, if voters only mistake opportunism for corruption in the pool of challengers, then bias makes them pessimistic ($\pi_0 = \pi_0^u \Rightarrow \eta^* < \eta^u$). In the opposite limit case, if voters only mistake opportunism for honesty in the pool of challengers, then bias makes them optimistic ($\pi_0 - \pi_0^u = \kappa^u - \kappa \Rightarrow \eta^* > \eta^u$).

Proposition 13 establishes that, intuitively, the FAE is more likely to induce mistaken demand for dictatorship by politicians who seem potentially Honest than by those whose dishonesty has transpired. Here, as in Proposition 11, misperception of the pool of politicians and of the option value of screening reinforce each other when perception of opportunism change. Biased voters who underestimate opportunism overestimate the probability that an incumbent who did not extract rents must be Honest, and simultaneously underestimate the benefit of the challenger's first-term electoral incentives.

On the other hand, misperceptions of honesty have ambiguous effects, because they lead voters to overestimate both the incumbent and his challenger. Either dimension of optimism may prevail. Thus, $\Xi_0 > 0$ if and only if

$$\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right] > \left[1 + \left(1 - \frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\left(1 - \pi_1^u\right). \tag{33}$$

If this condition is satisfied, mistakes in demand for dictatorship follow a similar pattern whether the incumbent did or did not extract rents. The FAE makes voters too keen on dictatorship unless it sufficiently raises perceived honesty as it lowers perceived opportunism. If perceptions of honesty are unaffected, voters are too keen on dictatorship by all politicians. There is a non-empty range of increases in perceived honesty such that voters are too keen on dictatorship by incumbents who appear Honest, but not by those who demonstrated dishonesty. If Equation (33) does not hold, in fact, the FAE necessarily increases demand for dictatorship by politicians who refrained from rent extraction—while it may well reduce it for politicians who extracted rents.

Figure 12 illustrates the impact of the FAE on the demand for dictatorship. This figure reflects a simulation with the same parameter values as before ($\delta = 1$, $\gamma_{i,t} \sim U[0,1]$, $\hat{r} = \hat{\eta} = 2\upsilon$). We have assumed that 50 percent of the politicians are Opportunists. Non-Opportunists are split equally between always Corrupt and always Honest leaders. The FAE causes voters to underestimate the share of politicians who are opportunists by half, and consequently to increase the perceived share of the other two groups by half each.

The graph shows five lines, which capture the true and perceived expected value to voters of electing a challenger, turning a leader who has engaged in rent extraction into a dictator, and allowing a leader who has not engaged in rent extraction to become a dictator. The expected values for the incumbent are naturally a function of the incumbent's ability level. The expected values for the challenger are independent of incumbent ability. There is only one expected value for politicians who have extracted rents, since the FAE beliefs about the returns from re-electing such leaders are exactly correct.

If the EV for an incumber exceeds the EV for a challenger, then the voters will support a bid for dictatorship from that incumbent. The two expected values for challengers are almost identical, although the perceived expected value lies slightly above the true expected value. While it is possible that the FAE significantly increases the appeal of challengers, under these parameter values it does not. Naturally, the FAE also fails to impact the returns to empowering a corrupt incumbent.

However, the FAE does significantly increase the returns associated with turning a non-rent extracting incumbent into a dictator. The FAE causes the perceived returns to empowering the non-extractors to increase sizably. This boost means that the minimum ability cutoff for approving a dictator is much lower with the FAE, and the demand for dictatorship has gone up.

Figure 13 increases the share of leaders who are Opportunists to 90 percent. The non-Opportunists are again equally split between always Honest and always Corrupt politicians. The other parameters are the same. Since the FAE causes voters to believe that the share of Opportunists is fifty percent smaller, in this case, that causes the perceived share of always Honest and always Corrupt leaders to increase by 450 percent each. In this case, the perceived expected value of the challenger is more noticeably higher than the true expected value of the challenger, reflecting the fact that FAE voters think that a new leader has a good chance of being a truly Honest politician. These parameter values mean that few politicians are truly Honest.

But the impact of the FAE on the expected value of empowering an incumbent who refrains from rent extraction is far larger. The true expected value of empowering a non-rent extractor is low, and close to the returns to empowering a rent-extractor. The overwhelming majority of the leaders who don't extract are Opportunists, and they will behave just like the rent extractors if they are turned into dictators.

However, FAE voters believe that the politicians who refrain from rent extraction are quite likely to be truly Honest. Consequently, they perceive the returns from turning such politicians into dictators to be quite high and the ability cutoff for empowering them is much lower. The FAE promotes the demand for dictatorship by causing voters to think that leaders who were honest only because of incentives are actually intrinsically honest, and hence can be trusted with unfettered power.

5 Implications of the Two Models Together

In this section, we discuss the normative and positive implications of both models together. We believe that both models provide a common view of how the FAE should change our expectations about the functioning of democracies, and we explore that first. We then discuss possible empirical tests of the two models.

5.1 Normative Implications of the FAE for Democracy

Ashworth and Bueno de Mesquita (2014) emphasize that semi-rationality can have positive effects on democracy, primarily by changing the equilibrium behavior of politicians. Our work confirms that conclusion. In our signaling model, the FAE typically reduces rent-seeking behavior by politicians. Voters attribute good behavior by politicians to permanent characteristics, not temporary incentives, and consequently good behavior carries more clout in the voting booth, which in turn creates stronger incentives for good behavior.

A variant of this effect also occurs in the signal-jamming model, when there are just two government products, a public service which is a function of exogenous shocks and public spending, and a second attribute of the politician, such as personality or ideology, that is independent of public spending. As long as the role of exogenous shock is lower in generating the politician's attribute, the FAE will lead to less rent extraction in the signal-jamming model as well. This effect generalizes to multiple services, as long as the services with most signal-to-noise also have higher returns to spending.

The signal-jamming model provides a more nuanced picture when there are multiple public services. The effect of the FAE is to push spending towards areas that have more true exogenous variation, since the FAE causes voters to put too much value on performance in those areas. This switch to higher signal-to-noise services can be benign, and is benign if services only differ in their signal-to-noise ratio. In that case, voter welfare is enhanced if politicians spread cash more evenly rather than favoring those services that have a low signal-to-noise ratio for electoral reasons. But if those services with high signal-to-noise ratios also have low returns from spending, then the FAE's impact on political behavior can harm welfare.

But while the FAE can have benign impacts on politicians' behavior, these benefits of the error must be placed against two costs: worse selection of politicians and bad institutional design. We suspect that the welfare losses due to bad screening of politicians is the less problematic issue. In the signal-jamming model, inferior screening means that voters are more likely to elect politicians who have gotten lucky, as in Wolfers (2007), and look like they have skills in areas where luck plays a disproportionate role. For example, this may mean that American voters elect Presidents to manage the economy, which many economists believe is largely driven by events outside the Presidents' control, instead of Presidents who specialize in foreign affairs, which seem much more squarely under presidential control. Since FAE voters attribute economic success to good political leadership, the model predicts that recent economic events will shape electoral outcomes, as it appears they do (Fair 1978).

In the signaling model, the FAE's effects on the selection of politicians are more complicated, but the most intuitive and for many parameter values the largest effects are that the FAE increases the re-election probability of politicians who have behaved well and reduces the re-election probability of politicians who have behaved poorly. Politicians who have behaved well are thought to be truly honest, rather than just opportunists who respond to incentives. Consequently, the FAE leads to too much trust in such leaders. The FAE means that politicians who have behaved badly, however, are thought to be permanently dishonest and the FAE can also increase the value of getting a new leader. Consequently, it can lead to too much toughness of politicians who have behaved badly.

We cannot tell whether the positive effects of the FAE on political behavior offset the negative effects on politician selection. Consequently, we cannot conclude that it would be socially beneficial to try to reduce the FAE in an election, if that could even be done. However, the impact of the FAE on institutional design seems more obviously negative, especially since there are no obvious offsetting behavioral benefits to bad institutional choices.

In the signal-jamming model, we introduced an institution that would provide more information to voters about the true contribution of the politician to public service provision. This institution could be a free press or more open political competition. The FAE means that voters under-estimate the value of such institutions, because they believe that what they already observe reveals more than it does about the politician's own contribution. As they believe that they already know the politician's worth without added help, they are consequently willing to pay less for acquiring more information about the politician's contributions.

This result may explain the limited protest about state subversion of the press in Turkey and Russia. In the wake of the failed 2016 coup d'état in Turkey, President Erdogan clamped down on the press, closing more than 100 media outlets and arresting numerous journalists. While western media observers howled, domestic opposition was quite limited. Similarly, Russian objections to President Putin's domination of the press also seem muted.

In the signaling model, the FAE led to a heightened demand for dictatorship for leaders who have behaved well in their first term of office. This result would generalize to settings in which voters have had alternative information about the behavior of leaders, including past military service, business careers or even behavior as a political candidate. The FAE would lead those voters to overestimate the extent to which that behavior indicates permanent individual strengths and consequently a greater willingness to assign permanent power to this person.

The FAE can even lead to too much demand for dictatorship in settings where the identity of the leader is unclear. Consider an extreme case in which everyone is actually an opportunist, but the FAE leads people to believe that everyone is either always good or always bad. A standard electoral system could lead to good incentives for the opportunists, that the extreme FAE voters would not foresee. A dictatorship creates no such incentives, but the FAE voter wouldn't care and might well prefer the advantage of durable leadership over the vicissitudes of democracy.

We believe that the potential risks of the FAE for institutional design suggests that careful limitations on institutional change could be welfare improving. Slow-moving constitutions that require lengthy debate before change would seem to provide a check on the errors created by the FAE. If professional training can reduce the FAE in politics, then there is also a case for delegating institutional design to constitutional conventions, whose designs are then ratified by the electorate.

5.2 Positive Implications

We separate our positive implications into three settings: the lab, voter behavior and politicians' behavior. In all cases, the great difficulty with empirical work on the FAE is the lack of obvious measures about what people suffer more or less from the error. A second difficulty is that suffering from the error is likely to be correlated with a host of other behaviors.

Testing the model is easiest in the laboratory. For example, a two-part signaling model in which one actor is endowed with an quantity of money that reflects a permanent component, and a temporary component. That actor is allowed to allocate some fraction of that amount to himself and the rest goes to a second player. The second player then decides to keep this partner for the second period, or to replace him with a new one-period actor. If retained, the first player can then decide about whether to cheat the player during the first period.

This game can be played with totally free actions in both periods, or conversely, actions can be artificially restricted during the first period. The experiment can also elicit the second player's beliefs about actions during the second period, if the first-period actor is retained. The model predicts that the second player will overestimate the tendency of the first player to repeat good behavior during the second period. Similarly, variants of signal-jamming model can be tested in the lab, and so can the prediction about the demand for dictatorship.

However, such laboratory work would essentially be yet another test of the existence of the FAE, rather than a more tangible real world prediction of the model. Some empirical implications of the model are really tests of the basic assumptions rather than tests of the model's implications. For example, the work of Wolfers (2007) showing the politicians are rewarded for luck also tests the implication of the model that FAE voters will re-elect politicians who have gotten lucky in office.

A related implication is that voters are routinely disappointed by the behavior of term-limited politicians during their final terms in office. It is true, for example, that since World War II, Gallup Presidential Approval Ratings have averaged 58.3 percent during the first term of two term presidents and 47.6 percent during the second term, yet many factors impact such polls, including lame-duck effects that limit presidential power and mean reversion.[6] Moreover, US presidents are sharply constrained in many ways, so they may not be able to pursue their own objectives instead of popular policies during their second terms.

It would be particularly interesting to test whether voters that are more affected by the FAE were also more likely to re-elect politicians for luck. This could be done either with large-scale surveys that administer tests for FAE-like behavior and that also ask about voting behavior and opinions about incumbent politicians. Alternatively, if the FAE were more prevalent in voters with particular demographics, such as lower levels of education (as in Benjamin and Shapiro 2007), then the model could be tested by examining whether voters with these demographics were also more likely to vote based on luck or be disappointed with final-term incumbents.

Perhaps, the hardest tests of the model are the implications about how the FAE impacts entire political systems. Some implications of the model are directly about observables, such as politicians eschewing corruption when they face more competitive re-election, but these predictions are common to many political models. The novel predictions of our model concern the implications of the FAE, and the level of the FAE in the population as a whole is unknown and possibly unknowable.

Again, perhaps the strongest possibility is to establish correlates of the FAE in demographic data. The model would then predict that localities or states in which the population has demographics that predict a greater tendency to exhibit the FAE would generate a greater gap between first-term and final-term behavior in term-limited officials and more regular voter disappointment. Yet we accept that these tests may be far-fetched.

There is, however, one test of the model that does seem to be widely supported by anecdote and history. Voters afflicted with the FAE should respond by demanding new people, not new institutions. They should want to "throw the bums out," rather than to change the incentives facing political actors. To our eyes, this tendency to respond to endemic corruption by hoping that the next leader will have a better character, instead of systemic reform, is commonplace in many settings.

6 The FAE and Consumer Behavior

The bulk of this paper has focused on the FAE's application to voters beliefs about politicians. Yet the FAE may also impact the beliefs that voters have about other citizens' behavioral response to economic policies. Consequently, they may underestimate the behavioral response to a tax or a new piece of government spending. In this section, we briefly sketch the impact of such a misperception.

Consider, for example, a government policy which features a tax rate t on a particular activity a and per capita government spending x. Aggregate behavior, denoted $\bar{a}$, directly influences consumer utility through an externality and determines the total tax revenues and thus spending $x = t\bar{a}$.

[6] There are eight presidents who were re-elected in office since 1945. The average gap between first and second term approval ratings was 10.7 percent. This gap drops to seven percent when Johnson and Truman are excluded, which is reasonable since they could plausibly run again.

The FAE takes the form of a mistaken belief that $d\bar{a}/dt$ is lower than in reality, because behavior is determined by type, not external circumstances. If the true relationship is $d\bar{a}/dt$, voters believe that the relationship is $(1-\beta)\, d\bar{a}/dt$, with $\beta \in [0,1]$. The behavior may have direct consequences for utility and may also impact consumer spending (or income), so that the price (or opportunity cost) of activity a net of taxes equals pa.

Individual welfare is $u = U(c, a, x, \bar{a})$, for consumption $c = y - (p+t)\, a$. We assume a well-behaved utility function that is twice continuously differentiable and concave in each argument. Moreover, we assume sufficient concavity and sufficiently small cross-partial derivatives to ensure a unique, interior maximum.

We will focus our discussion on the simple case in which agents are truly homogeneous, so that in equilibrium $\bar{a} = a$ and $d\bar{a}/dt = da/dt$. In this context, the FAE assumption means that individuals correctly anticipate their own response, but believe that the average market response is only a fraction $(1-\beta)$ of it.

6.1 Political-Behavioral Equilibrium

We define a political-behavioral equilibrium by the following two properties.

1. Individuals choose their own behavior to maximize utility, taking the tax rate, the government subsidy, and the behavior of other agents as given.

2. The tax rates maximizes perceived welfare for the representative agent, who observes equilibrium quantities, and believes that the aggregate response to a change in taxes will equal $(1-\beta)$ times her own response.

The first-order condition for the consumers' utility maximization problem is:

$$U_a = (p+t)\, U_c, \tag{34}$$

omitting arguments for the sake of brevity. For a well-behaved utility function, this optimality condition implicitly defines individual action as a function of market conditions: $a = A(y, p+t, x, \bar{a})$. Moreover, any tax rate t induces a unique, interior market equilibrium once we rule out extreme complementarity between individual and average activity.[7]

The FAE causes voters to underestimate the responsiveness of average activity $\bar{a}$ to changing conditions. Thus, the first-order condition for tax setting is:

$$\bar{a}U_x + (U_{\bar{a}} + tU_x)(1-\beta)\frac{d\bar{a}}{dt} = aU_c. \tag{35}$$

Again, for a well-behaved utility function this optimality condition defines a unique, interior tax rate that voters mistakenly believe to be welfare maximizing.

The FAE will lead to excessive taxation when $(U_{\bar{a}} + tU_x)\, d\bar{a}/dt > 0$. This is the case when taxes fund general government spending and negatively impact labor supply. Taxation will instead be insufficient when $(U_{\bar{a}} + tU_x)\, d\bar{a}/dt < 0$. This is the case when Pigovian taxes deter behavior that creates negative externalities. We now use this general framework to provide results for a few common public-policy problems.

[7] Formally, so long as $tA_x + A_{\bar{a}} < 1$ there is a unique, interior equilibrium in which the true behavioral response is $d\bar{a}/dt = (A_p + \bar{a}A_x)/(1 - tA_x - A_{\bar{a}})$. This is assured if all cross-partial derivatives of U are approximately nil.

6.2 The FAE and the Size of Government

In this case, we assume that a represents labor supply, so $U_a < 0$ and $p < -w$ where $w > 0$ is the pre-tax wage. We assume that $U_{\bar{a}} = 0$, so voters have no interest in the working hours of others. Moreover, for simplicity, we assume quasi-linearity and separability so that utility equals:

$$u = y + (w - t)\, a - C\,(a) + V\,(x)\,, \tag{36}$$

for a convex cost of effort C and a concave value of public goods V.

The first-order condition for labor supply is:

$$w - t = C'\,(a) \text{ such that } \frac{da}{dt} = -\frac{1}{C''\,(a)} < 0. \tag{37}$$

The first-order condition for perceived welfare-maximizing taxation is:

$$V'\,(t\bar{a}) \left[\bar{a} - \frac{(1-\beta)\, t}{C''\,(\bar{a})}\right] = a, \tag{38}$$

and we assume that V is sufficiently concave to satisfy the second-order condition for a maximum. It is then immediate that the FAE increases taxation ($\partial t / \partial \beta > 0$).

When applied to tax policy, a tendency to underestimate behavioral responses means that voters fail to recognize that higher taxes will deter working in others. Consequently, they will underestimate the distortive impact of taxes and choose tax rates that are too high. We suspect that this effect may have been important in the adoption of high tax regimes globally, prior to 1970, and may still be important in European countries, where the tendency to attribute high earnings to luck is particularly strong (Alesina and Glaeser 2004).

In the US, since the late 1970s and the promulgation of ideas like the Laffer curve, American voters have long been deluged with messages arguing that tax rates do indeed reduce effort and higher taxes may even reduce total tax payments. It is possible that these messages have reduced, or even eliminated the FAE among American voters in this respect. It is also possible that anti-tax voters do not believe that government spending x will increase with the tax rate t, which can be construed as a different variant of the FAE. Naturally, a failure to link tax rates with desirable government spending can also be explained by a low assessment of the value of government spending or a belief that higher debt levels will not lead to higher taxes in the future.

6.3 The FAE, Pigouvian Taxes and Quantity Regulation

In this case we assume that the activity a yields private benefits but generates a public cost ($U_a > 0 > U_{\bar{a}}$). Taxation serves only to deter this negative externality, and tax revenues are simply rebated to voters after a fraction λ is absorbed by the cost of operating the system—which has been quite considerable in cases such as London's congestion charge. We assume again quasi-linearity and separability so that utility equals:

$$u = y - (p + t)\, a + (1 - \lambda)\, x + B\,(a) - C\,(\bar{a})\,, \tag{39}$$

for a concave private benefit B and a convex social cost C.

In this case, the first-order condition for activity is:

$$B'(a) = p + t \text{ such that } \frac{da}{dt} = \frac{1}{B''(a)} < 0. \tag{40}$$

The first-order condition for perceived welfare-maximizing taxation is:

$$(1-\lambda)\left[\bar{a} + \frac{(1-\beta)t}{B''(\bar{a})}\right] - (1-\beta)\frac{C'(\bar{a})}{B''(\bar{a})} = a, \tag{41}$$

and we assume that C is sufficiently convex to satisfy the second-order condition for a maximum.

When taxation is costless ($\lambda = 0$) the FAE does not impact the optimal level of the Pigouvian tax, which remains $t = C'(\bar{a})$. The classic Pigouvian optimum does not depend on a correct understanding of the behavioral response to the tax. This result changes, however, as soon as taxation involves any form of waste ($\lambda > 0$). Then, recalling that $\bar{a} = a$ in equilibrium, the optimality condition can be rearranged:

$$(1-\beta)\frac{(1-\lambda)t - C'(\bar{a})}{B''(\bar{a})} = \lambda\bar{a} > 0, \tag{42}$$

showing that the FAE reduces taxation ($\partial t/\partial\beta < 0$). Since the FAE leads to underestimating the impact of the Pigouvian tax on behavior, it will also lead to a tax that is too small.

This result seems relevant for the relative lack of popularity of Pigouvian policies, such as congestion pricing. As long as these policies involve some waste, they will not fully internalize the social costs of adverse action. The presence of the FAE means that voters don't believe that these taxes will do much to reduce the driving of others. The costs are fully anticipated, but the benefits are not, and consequently Pigouvian taxes will not be fully implemented.

The experience of Stockholm seems particularly relevant. Initially, the congestion tax was unpopular, but after a trial run in the first half of 2006, a referendum was held. The experience of the tax appears to have changed voters' minds and the congestion tax was subsequently approved by a margin of 52.5 to 47.5 percent. This switch in attitude is compatible with an initial belief that the congestion charge would do little to impact behavior, which changed after the experience of the tax.

Regulations that restrict $\bar{a}$ directly will not be subject to the FAE, and so we should expect to see more quantity restrictions in practice than economic theory would typically suggest. Indeed, the presence of clumsy quantity restrictions, including CAFE standards on automobiles, even in cases where a Pigouvian tax would seem to be more efficient, is a significant policy puzzle. The FAE can explain why such restrictions have more appeal with voters than taxes, since with the FAE voters underestimate the impact that these taxes will have on behavior.

6.4 The FAE and Infrastructure

A final area of policy is the construction of infrastructure. Duranton and Turner (2011) have documented a "fundamental law" of highway traffic: vehicle miles travelled increase roughly one-for-one with highway miles built. This finding supports the earlier Downes' Law (1962) that peak-hour traffic speeds are independent of the amount of road space. Both laws argue that the benefit of infrastructure is significantly muted by an increase in the use of the infrastructure.

In this case we assume that the benefits of infrastructure use a are reduced by congestion, which is a function of the ration of total usage $\bar{a}$ to infrastructure spending s. Thus, utility equals:

$$u = y - s + B(a) - aC\left(\frac{\bar{a}}{s}\right), \tag{43}$$

for a concave benefit B and a convex congestion cost C.[8]

The first-order condition for individual infrastructure usage is

$$B'(a) = C\left(\frac{\bar{a}}{s}\right), \tag{44}$$

and when $B(a) = ba$ this implies that the ratio $\bar{a}/s$ is independent of s, so the fundamental law holds exactly. More generally, individual usage obeys $a = A(s, \bar{a})$ with derivatives:

$$A_s = -\frac{\bar{a}}{s^2}\frac{C'(\bar{a}/s)}{B''(a)} > 0 > A_{\bar{a}} = \frac{1}{s}\frac{C'(\bar{a}/s)}{B''(a)}. \tag{45}$$

Thus, any amount of infrastructure determines a unique market equilibrium $\bar{a} = A(s, \bar{a})$ in which the true sensitivity of aggregate usage to the amount of infrastructure is

$$\frac{d\bar{a}}{ds} = \frac{A_s}{1 - A_{\bar{a}}} = \frac{\bar{a}}{s}\frac{C'(\bar{a}/s)}{C'(\bar{a}/s) - sB''(\bar{a})} > 0, \tag{46}$$

but voters misperceive it as $(1 - \beta)$ times this amount.

The first-order condition for spending on infrastructure is then:

$$\frac{a}{s}C'\left(\frac{\bar{a}}{s}\right)\frac{\bar{a}}{s}\frac{\beta C'(\bar{a}/s) - sB''(\bar{a})}{C'(\bar{a}/s) - sB''(\bar{a})} = 1, \tag{47}$$

and we assume again that the second-order condition for a maximum is satisfied. It is then immediate that the FAE leads to too much voter desire for infrastructure ($\partial s/\partial\beta > 0$).[9]

In this case, the behavioral response to spending is more driving, which in turn lowers voter welfare. Voters fail to anticipate the full extent of this response and this leaves them overly enthusiastic about more road building. This model predicts that voters will be regularly surprised that new roads or bridges do not materially shorten their daily commutes.

[8] The model remains equivalent to the general case presented above. If we defined the "congestion tax" $t = C(\bar{a}/s)$ paid by each individual, we could write $u = y + B(a) - ta - \bar{a}/C^{-1}(x/\bar{a})$. The last term is the value of the public goods (lower income taxes) that result from allowing more congestion: a monotone increasing and concave function of the congestion ($x = t\bar{a}$) suffered by others, as well as a monotone decreasing and concave function of their usage.

[9] Analogously, if we rewrote the model in terms of a "congestion tax," we could derive the first-order condition for consumption: $B' = t$, such that $da/dt = 1/B'' < 0$; and the first-order condition for policy: $\bar{a}/\left[C'\left(C^{-1}\right)^2\right] - (1-\beta)/\left(B''C^{-1}\right) = a$, such that by the second-order condition for a maximum $\partial t/\partial\beta < 0$. The FAE induces voters to demand too low a "congestion tax," which is to say too much infrastructure spending.

Hence, this model can perhaps explain why voters have often been more supportive of new infrastructure projects than transport economists.

At this point, we are not making claims that the FAE is the key driver of voting behavior towards new roads, or Pigouvian taxes, or the overall level of government spending. The models do suggest implications that are broadly compatible with a voting public that places too much faith in new roads and too little faith in the pricing of those roads. We hope future work will more thoroughly test the implications of these models.

7 Conclusion

In this paper, we have accepted a claim that is common in the psychology literature: individuals underestimate the importance of circumstance and overstate the importance of enduring personal characteristics. Our acceptance of this claim was supported by the work of Wolfers (2007) and others, showing that voters rewarded politicians for luck. We then applied this assumption to two standard models in the political science literature: the signal-jamming model of Alesina and Tabellini (2008) and the signaling model of Besley (2007).

The models revealed similar conclusions. As in Ashworth and Bueno de Mesquita (2014), voter irrationality actually improved the behavior of politicians in many of the most natural settings. When voters think that good behavior in the first term of a politician is more likely to reflect a permanent character trait rather than the impact of incentives, then opportunistic politicians are more likely to behave well. When voters think that the heterogeneity in outcomes reflects personal ability more than luck, politicians will try harder to boost those outcomes.

Yet these enhanced incentives come at a cost. The FAE naturally leads voters to make mistakes about re-election, such as overweighting skills in areas where politicians actually have little impact relative to skills in areas where politicians really make a difference. Like many economists, we believe that American Presidents have far more ability to impact foreign affairs than to impact the economy. The FAE might explain why voters do seem to favor politicians who claim that they can manage the economy.

Most troublingly, the FAE will lead to mistakes about institutional design and reform. In response to corruption, voters will simply choose to eject corrupt leaders, failing to anticipate that their replacements will face the same incentives and be just as corrupt. With the FAE, voters will fail to see the benefits of changing incentives to produce better leadership outcomes.

The FAE will also imply too little demand for better information and too much demand for dictatorship. Better information might help voters make better election decisions by helping them to distinguish signal from noise. Voters with the FAE believe that the signal-to-noise ratio is intrinsically high, and consequently they have little need for alternative sources of analysis such as a free press.

Voters with the FAE will also favor endowing well-performing leaders with too much power, failing to recognize that their behavior will change when they no longer face electoral incentives. Similarly, the FAE could also explain why voters see little value in checks and balances, but instead seek merely an intrinsically "good" leader who will serve their interests. Too much faith in the person, relative to the system, can create a tragic lack of interest in providing better political institutions.

In this paper, we also produced some speculative thoughts on the FAE as applied to policies that directly impact consumers. This work suggested that voters would be too enthusiastic about spending on infrastructure, since the fail to anticipate the behavioral response, and too hostile to demand management policies, such as congestion pricing. They may also fail to anticipate the behavioral consequences of higher taxes or subsidies.

References

[1] Alesina, Alberto, and Guido Tabellini. 2007. Bureaucrats or Politicians? Part I: A Single Policy Task. *American Economic Review* 97 (1): 169–179.

[2] Alesina, Alberto, and Guido Tabellini. 2008. Bureaucrats or Politicians? Part II: Multiple Policy Tasks. *Journal of Public Economics* 92 (3): 426–447.

[3] Ashworth, Scott, and Ethan Bueno de Mesquita. 2014. Is Voter Competence Good for Voters? Information, Rationality, and Democratic Performance. *American Political Science Review* 108 (3): 565–587.

[4] Banks, Jeffrey S., and Joel Sobel. 1987. Equilibrium Selection in Signaling Games. *Econometrica* 55 (3): 647-661.

[5] Bertrand, Marianne, and Sendhil Mullainathan. 2001. Are CEOs Rewarded for Luck? The Ones without Principals Are. *Quarterly Journal of Economics* 116 (3): 901–932.

[6] Besley, Timothy. 2007. *Principled Agents? The Political Economy of Good Government.* Oxford: Oxford University Press.

[7] Bierbrauer, Günter. 1974. Attribution and Perspective: Effect of Time Set and Role on Interpersonal Inference. Doctoral dissertation, Stanford University.

[8] Boffa, Federico, Amedeo Piolatto, and Giacomo A. M. Ponzetto. 2016. Political Centralization and Government Accountability. *Quarterly Journal of Economics* 131 (1): 381–422.

[9] Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics* 119 (3): 861–898.

[10] Dal Bó, Ernesto, Pedro Dal Bó, and Erik Eyster. 2017. The Demand for Bad Policy When Voters Underappreciate Equilibrium Effects. *Review of Economic Studies*, online access https://doi.org/10.1093/restud/rdx031.

[11] Eyster, Erik, and Matthew Rabin. 2005. Cursed Equilibrium. *Econometrica* 73 (5): 1623–1672.

[12] Fair, Ray C. 1978. The Effect of Economic Events on Votes for President. *Review of Economics and Statistics* 60 (2): 159–173.

[13] Gilbert, Daniel T., and Patrick S. Malone. 1995. The Correspondence Bias. Psychological Bulletin 117 (1): 21–38.

[14] Grossman, Gene M., and Elhanan Helpman. 2001. *Special Interest Politics.* Cambridge, MA: MIT Press.

[15] Jones, Edward E., and Victor A. Harris. 1967. The Attribution of Attitudes. *Journal of Experimental Social Psychology* 3 (1): 1–24.

[16] Jones, Edward E., and Richard E. Nisbett. 1971. *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior.* New York, NY: General Learning Press.

[17] Malle, Bertram F. 2006. The Actor-Observer Asymmetry in Attribution: A (Surprising) Meta-Analysis. *Psychological Bulletin* 132 (6): 895–919.

[18] McArthur, Leslie A. 1972. The How and What of Why: Some Determinants and Consequences of Causal Attribution. *Journal of Personality and Social Psychology* 22 (2): 171–193.

[19] Milgram, Stanley. 1963. Behavioral Study of Obedience. *Journal of Abnormal and Social Psychology* 67 (4): 371–378.

[20] Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy.* Cambridge, MA: MIT Press.

[21] Rogoff, Kenneth. 1990. Equilibrium Political Budget Cycles. *American Economic Review* 80 (1): 21–36.

[22] Ross, Lee. 1977. The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process. *Advances in Experimental Social Psychology* 10: 173–220.

[23] Wolfers, Justin. 2007. Are Voters Rational? Evidence from Gubernatorial Elections. Mimeo, University of Michigan.

A Mathematical Appendix

A.1. Proof of Proposition 1 and Corollary 1

The incumbent is re-elected with probability

$$p(\mathbf{x}_t) = \Phi\left(\frac{\sum_{g=1}^{G}\alpha_g\left[1-(1-\beta)\nu_g\right]\rho_g\left(\ln x_{g,t}-\ln\bar{x}_g\right)}{\sqrt{\sum_{g=1}^{G}\left\{\alpha_g\left[1-(1-\beta)\nu_g\right]\sigma_g\right\}^2}}\right), \tag{A1}$$

where Φ denotes the cumulative distribution function of the standard normal distribution.

Denoting by R the incumbent's value of re-election, he chooses expenditure

$$\mathbf{x}_t = \arg\max_{\mathbf{x}}\left\{b-\sum_{g=1}^{G}x_{g,t}+Rp(\mathbf{x}_t)\right\}. \tag{A2}$$

The first-order conditions of this problem are:

$$\begin{aligned} x_{g,t} &= \alpha_g\left[1-(1-\beta)\nu_g\right]\rho_g \\ &\times\frac{R}{\sqrt{\sum_{j=1}^{G}\left\{\alpha_j\left[1-(1-\beta)\nu_j\right]\sigma_j\right\}^2}}\phi\left(\frac{\sum_{j=1}^{G}\alpha_j\left[1-(1-\beta)\nu_j\right]\rho_j\left(\ln x_{j,t}-\ln\bar{x}_j\right)}{\sqrt{\sum_{j=1}^{G}\left\{\alpha_j\left[1-(1-\beta)\nu_j\right]\sigma_j\right\}^2}}\right), \end{aligned} \tag{A3}$$

where ϕ denotes the probability density function of the standard normal distribution.

In a rational expectations equilibrium the budget allocation is

$$\bar{x}_g = x_{g,t} = \frac{\alpha_g\left[1-(1-\beta)\nu_g\right]\rho_g R}{\sqrt{2\pi\sum_{j=1}^{G}\left\{\alpha_j\left[1-(1-\beta)\nu_j\right]\sigma_j\right\}^2}}, \tag{A4}$$

with rent extraction

$$r = b-\frac{R\sum_{g=1}^{G}\alpha_g\left[1-(1-\beta)\nu_g\right]\rho_g}{\sqrt{2\pi\sum_{g=1}^{G}\left\{\alpha_g\left[1-(1-\beta)\nu_g\right]\sigma_g\right\}^2}}. \tag{A5}$$

The incumbent is re-elected if and only if

$$\sum_{g=1}^{G}\alpha_g\left[1-(1-\beta)\nu_g\right]\left(\eta_{g,t}^{i_t}+\varepsilon_{g,t}\right)\geq 0, \tag{A6}$$

so the probability of re-election each period is $p(\bar{\mathbf{x}}) = 1/2$. The value of re-election is then

$$R = \delta\sum_{t=0}^{\infty}\left(\frac{\delta}{2}\right)^t r = \frac{2\delta}{2-\delta}r, \tag{A7}$$

so the stationary level of rent extraction is

$$r = b\left[1+\sqrt{\frac{2}{\pi}}\frac{\delta}{2-\delta}\frac{\sum_{g=1}^{G}\alpha_g\left[1-(1-\beta)\nu_g\right]\rho_g}{\sqrt{\sum_{g=1}^{G}\left\{\alpha_g\left[1-(1-\beta)\nu_g\right]\sigma_g\right\}^2}}\right]^{-1}. \tag{A8}$$

Corollary 1 follows immediately by setting $\nu_g = \nu$ for all g and simplifying the formulae.

A.2. Proof of Proposition 2

The true expected ability of a ruling politician, conditional on his being in office, is

$$\mathbb{E}\eta_{g,t-1}^{i_t} = \int_{-\infty}^{\infty} \frac{\eta}{\sigma_g\sqrt{1-\nu_g}}\phi\left(\frac{\eta}{\sigma_g\sqrt{1-\nu_g}}\right)\Phi\left(\frac{\eta}{\Sigma_g}\right)d\eta \tag{A9}$$

for

$$\Sigma_g \equiv \sqrt{\frac{\sum_{j=1}^{G}\left\{\alpha_j\left[1-(1-\beta)\nu_j\right]\sigma_j\right\}^2}{\left\{\alpha_g\left[1-(1-\beta)\nu_g\right]\right\}^2} - (1-\nu_g)\,\sigma_g^2}, \tag{A10}$$

such that

$$\begin{aligned}\frac{\partial\mathbb{E}\eta_{g,t-1}^{i_t}}{\partial\beta} &= -\frac{1}{\Sigma_g}\frac{\partial\Sigma_g}{\partial\beta}\int_{-\infty}^{\infty}\frac{\eta^2}{\sigma_g\sqrt{1-\nu_g}}\phi\left(\frac{\eta}{\sigma_g\sqrt{1-\nu_g}}\right)\frac{1}{\Sigma_g}\phi\left(\frac{\eta}{\Sigma_g}\right)d\eta\\ &= -\frac{(1-\nu_g)\,\sigma_g^2\Sigma_g}{\sqrt{2\pi\left[(1-\nu_g)\,\sigma_g^2+\Sigma_g^2\right]^3}}\frac{\partial\Sigma_g}{\partial\beta}\\ &= \frac{\alpha_g(1-\nu_g)\,\sigma_g^2}{\sqrt{2\pi\left(\sum_{j=1}^{G}\left\{\alpha_j\left[1-(1-\beta)\nu_j\right]\sigma_j\right\}^2\right)^3}}\sum_{j=1}^{G}(\nu_g-\nu_j)\,\alpha_j^2\left[1-(1-\beta)\nu_j\right]\sigma_j^2.\end{aligned} \tag{A11}$$

The impact of distorted screening on welfare is captured by:

$$\sum_{g=1}^{G}\alpha_g\frac{\partial\mathbb{E}\eta_{g,t-1}^{i_t}}{\partial\beta} = \beta\frac{\left(\sum_{g=1}^{G}\alpha_g^2\nu_g\sigma_g^2\right)^2 - \sum_{g=1}^{G}(\alpha_g\sigma_g)^2\sum_{g=1}^{G}(\alpha_g\nu_g\sigma_g)^2}{\sqrt{2\pi\left(\sum_{g=1}^{G}\left\{\alpha_g\left[1-(1-\beta)\nu_g\right]\sigma_g\right\}^2\right)^3}} \le 0.$$

To see that the numerator is always negative, let

$$f_g \equiv \frac{(\alpha_g\sigma_g)^2}{\sum_{j=1}^{G}(\alpha_j\sigma_j)^2} \ge 0 \text{ such that } \sum_{g=1}^{G} f_g = 1. \tag{A12}$$

Then the numerator equals

$$\left[\sum_{g=1}^{G}(\alpha_g\sigma_g)^2\right]^2\left[\left(\sum_{g=1}^{G}\nu_g f_g\right)^2 - \sum_{g=1}^{G}\nu_g^2 f_g\right] \tag{A13}$$

and f_g is well defined as the probability mass function of a discrete random variable ν with G distinct values ν_g, so

$$\sum_{g=1}^{G}\nu_g^2 f_g - \left(\sum_{g=1}^{G}\nu_g f_g\right)^2 = \operatorname{Var}(\nu) > 0. \tag{A14}$$

A.3. Proof of Proposition 3

The equilibrium share of productive expenditure devoted to public service g is

$$\xi_g \equiv \frac{\alpha_g \left[1 - (1-\beta)\nu_g\right] \rho_g}{\sum_{j=1}^{G} \alpha_j \rho_j \left[1 - (1-\beta)\nu_j\right]}, \tag{A15}$$

such that

$$\frac{\partial \xi_g}{\partial \beta} = \frac{\alpha_g \rho_g \sum_{j=1}^{G} (\nu_g - \nu_j) \alpha_j \rho_j}{\left\{\sum_{j=1}^{G} \alpha_j \rho_j \left[1 - (1-\beta)\nu_j\right]\right\}^2}. \tag{A16}$$

The impact of a distorted budget allocation on welfare is captured by:

$$\sum_{g=1}^{G} \alpha_g \rho_g \frac{\partial \ln \xi_g}{\partial \beta} = \sum_{g=1}^{G} \frac{\alpha_g \rho_g \nu_g}{1 - (1-\beta)\nu_g} - \frac{\sum_{g=1}^{G} \alpha_g \rho_g \sum_{g=1}^{G} \alpha_g \rho_g \nu_g}{\sum_{g=1}^{G} \alpha_g \rho_g \left[1 - (1-\beta)\nu_g\right]} \geq 0. \tag{A17}$$

To see that this is always positive, let

$$f_g \equiv \frac{\alpha_g \rho_g}{\sum_{j=1}^{G} \alpha_j \rho_j} \geq 0 \text{ such that } \sum_{g=1}^{G} f_g = 1. \tag{A18}$$

Then

$$\begin{aligned}\sum_{g=1}^{G} \alpha_g \rho_g \frac{\partial \ln \xi_g}{\partial \beta} = {} & \frac{\sum_{g=1}^{G} \alpha_g \rho_g}{\sum_{g=1}^{G} \left[1 - (1-\beta)\nu_g\right] f_g} \\ & \times \left[\sum_{g=1}^{G} \frac{\nu_g}{1 - (1-\beta)\nu_g} f_g \sum_{g=1}^{G} \left[1 - (1-\beta)\nu_g\right] f_g - \sum_{g=1}^{G} \nu_g f_g\right]\end{aligned} \tag{A19}$$

and f_g is well defined as the probability mass function of a discrete random variable ν with G distinct values ν_g, so

$$\begin{aligned}\sum_{g=1}^{G} \nu_g f_g - \sum_{g=1}^{G} \frac{\nu_g}{1 - (1-\beta)\nu_g} f_g \sum_{g=1}^{G} \left[1 - (1-\beta)\nu_g\right] f_g & \\ = \operatorname{Cov}\left(\frac{\nu}{1 - (1-\beta)\nu}, 1 - (1-\beta)\nu\right) & \leq 0.\end{aligned} \tag{A20}$$

A.4. Proof of Proposition 4

The impact of the FAE on rent extraction is captured by:

$$\begin{aligned}\frac{\partial r}{\partial \beta} & = \frac{(b-r)\,r}{b} \left(\frac{\sum_{g=1}^{G} \alpha_g^2 \left[1 - (1-\beta)\nu_g\right] \nu_g \sigma_g^2}{\sum_{g=1}^{G} \left\{\alpha_g \left[1 - (1-\beta)\nu_g\right] \sigma_g\right\}^2} - \frac{\sum_{g=1}^{G} \alpha_g \rho_g \nu_g}{\sum_{g=1}^{G} \alpha_g \rho_g \left[1 - (1-\beta)\nu_g\right]}\right) \\ & = \frac{(b-r)\,r}{b} \left(\frac{\sum_{g=1}^{G} \psi_g^2 \varsigma_g^2}{\sum_{g=1}^{G} \varsigma_g^2} - \frac{\sum_{g=1}^{G} \psi_g x_g}{\sum_{g=1}^{G} x_g}\right).\end{aligned} \tag{A21}$$

A.5. Proof of Corollary 2

If $\alpha_g = \alpha$ for all g then:

$$\frac{\partial r}{\partial \beta} = \frac{(b-r)\,r}{b} \frac{\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right] \sigma_g^2 \sum_{g=1}^{G} \rho_g}{\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right]^2 \sigma_g^2 \sum_{g=1}^{G} \rho_g \left[1-(1-\beta)\,\nu_g\right]} \times \left\{ \frac{\sum_{g=1}^{G} \nu_g \left[1-(1-\beta)\,\nu_g\right] \sigma_g^2}{\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right] \sigma_g^2} - \frac{\sum_{g=1}^{G} \rho_g \nu_g}{\sum_{g=1}^{G} \rho_g} \right\}. \quad \text{(A22)}$$

If moreover $\sigma_g^2 = \sigma^2$ for all g then:

$$\frac{\partial r}{\partial \beta} = -\frac{(b-r)\,r}{b} \frac{\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right] \sum_{g=1}^{G} \rho_g}{\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right]^2 \sum_{g=1}^{G} \rho_g \left[1-(1-\beta)\,\nu_g\right]} \times \left[(1-\beta) \frac{\frac{1}{G}\sum_{g=1}^{G} \nu_g^2 - \left(\frac{1}{G}\sum_{g=1}^{G} \nu_g\right)^2}{1-(1-\beta)\frac{1}{G}\sum_{g=1}^{G} \nu_g} + \frac{\frac{1}{G}\sum_{g=1}^{G} \rho_g \nu_g - \frac{1}{G}\sum_{g=1}^{G} \rho_g \frac{1}{G}\sum_{g=1}^{G} \nu_g}{\frac{1}{G}\sum_{g=1}^{G} \rho_g} \right]. \quad \text{(A23)}$$

If moreover $\rho_g = \rho$ for all g then:

$$\frac{\partial r}{\partial \beta} = -(1-\beta) \frac{(b-r)\,r}{b} \frac{\frac{1}{G}\sum_{g=1}^{G} \nu_g^2 - \left(\frac{1}{G}\sum_{g=1}^{G} \nu_g\right)^2}{\frac{1}{G}\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right]^2 \frac{1}{G}\sum_{g=1}^{G} \left[1-(1-\beta)\,\nu_g\right]} \leq 0. \quad \text{(A24)}$$

A.6. Proof of Corollary 3

If $G = 2$ and $\rho_2 = 0$ then

$$\frac{\partial r}{\partial \beta} = \frac{(b-r)\,r}{b} \frac{\alpha_2^2 \left[1-(1-\beta)\,\nu_2\right] \sigma_2^2}{\left[1-(1-\beta)\,\nu_1\right] \sum_{g=1}^{2} \left\{\alpha_g \left[1-(1-\beta)\,\nu_g\right] \sigma_g\right\}^2} (\nu_2 - \nu_1)\,. \quad \text{(A25)}$$

A.7. Equilibrium with Full Transparency

When exogenous conditions $\varepsilon_{g,t}$ are perfectly revealed, voter inference obeys Equation 11 and the incumbent is be re-elected if and only if

$$\sum_{g=1}^{G} \alpha_g \left[\eta_{g,t}^{i_t} + \rho_g \left(\ln x_{g,t} - \ln \bar{x}_g\right)\right] \geq 0. \quad \text{(A26)}$$

Voters' biased perception of the incumbent's probability of re-election is

$$\Phi\left(\frac{\sum_{g=1}^{G} \alpha_g \rho_g \left(\ln x_{g,t} - \ln \bar{x}_g\right)}{\sqrt{\sum_{g=1}^{G} \alpha_g^2 \left[1-(1-\beta)\,\nu_g\right] \sigma_g^2}} \right). \quad \text{(A27)}$$

If politicians share voters' biased beliefs, their optimization problem has first-order conditions:

$$x_{g,t} = \frac{\alpha_g \rho_g R}{\sqrt{\sum_{j=1}^{G} \alpha_j^2 \left[1-(1-\beta)\nu_j\right]\sigma_j^2}} \phi\left(\frac{\sum_{j=1}^{G} \alpha_j \rho_j \left(\ln x_{j,t} - \ln \bar{x}_j\right)}{\sqrt{\sum_{j=1}^{G} \alpha_j^2 \left[1-(1-\beta)\nu_j\right]\sigma_j^2}}\right). \tag{A28}$$

There is a unique stationary rational expectations equilibrium with the budget allocation

$$\bar{x}_g = x_{g,t} = \frac{\alpha_g \rho_g R}{\sqrt{2\pi \sum_{j=1}^{G} \alpha_j^2 \left[1-(1-\beta)\nu_j\right]\sigma_j^2}}, \tag{A29}$$

and rent extraction

$$r = b - \frac{R \sum_{g=1}^{G} \alpha_g \rho_g}{\sqrt{2\pi \sum_{g=1}^{G} \alpha_g^2 \left[1-(1-\beta)\nu_g\right]\sigma_g^2}}. \tag{A30}$$

The incumbent is re-elected if and only if $\sum_{g=1}^{G} \alpha_g \eta_{g,t}^{i_t} \geq 0$, hence with probability 1/2. The stationary level of rent extraction is

$$r = b\left[1 + \sqrt{\frac{2}{\pi}} \frac{\delta}{2-\delta} \frac{\sum_{g=1}^{G} \alpha_g \rho_g}{\sqrt{\sum_{g=1}^{G} \alpha_g^2 \left[1-(1-\beta)\nu_g\right]\sigma_g^2}}\right]^{-1}. \tag{A31}$$

A.8. Proof of Proposition 5

Denote by $A \equiv \sum_{g=1}^{G} \alpha_g \eta_{g,t-1}^{i_{t-1}}$ the welfare value of the ability of an incumbent running for re-election, and by $S \equiv \sum_{g=1}^{G} \alpha_g \left[1-(1-\beta)\nu_g\right]\left(\eta_{g,t-1}^{i_{t-1}} + \varepsilon_{g,t-1}\right)$ the voters' screening signal in the presence of noise. Then A and S are jointly normally distributed with means $\mu_A = \mu_S = 0$. Denote their variances by σ_A^2 and σ_S^2 and their correlation coefficient ρ.

Recall that the expectation of a truncated univariate normal distribution (Johnson, Kotz and Balakrishnan 1994, §13.10) is:

$$\mathbb{E}\left(A|A \geq a\right) = \mu_A + \sigma_A \frac{\phi\left(\frac{a-\mu_A}{\sigma_A}\right)}{1 - \Phi\left(\frac{a-\mu_A}{\sigma_A}\right)}; \tag{A32}$$

while the expectation of a truncated bivariate normal distribution (Kotz, Balakrishnan and Johnson 2000, §46.9) is:

$$\mathbb{E}\left(A|S \geq s\right) = \mu_A + \rho \frac{\sigma_A}{\sigma_S}\left[\mathbb{E}\left(S|S \geq s\right) - \mu_S\right]. \tag{A33}$$

Thus, the welfare value of ability under full transparency is:

$$\mathbb{E}\left(\sum_{g=1}^{G} \alpha_g \eta_{g,t-1}^{i_t^*}\right) = \Pr\left(A \geq 0\right)\mathbb{E}\left(A|A \geq 0\right) = \frac{\sigma_A}{\sqrt{2\pi}}, \tag{A34}$$

while the welfare value of ability with noise is:

$$\mathbb{E}\left(\sum_{g=1}^{G}\alpha_g\eta_{g,t-1}^{i_t}\right)=\Pr\left(S\geq 0\right)\mathbb{E}\left(A|S\geq 0\right)=\frac{\rho\sigma_A}{\sqrt{2\pi}}. \tag{A35}$$

Voters' mistake in assessing the value of transparency is thus:

$$\Delta=\frac{1}{\sqrt{2\pi}}\left[\left(1-\rho\right)\sigma_A-\left(1-\tilde{\rho}\right)\tilde{\sigma}_A\right]. \tag{A36}$$

The true variances of A and S are:

$$\sigma_A^2=\sum_{g=1}^{G}\alpha_g^2\left(1-\nu_g\right)\sigma_g^2 \text{ and } \sigma_S^2=\sum_{g=1}^{G}\alpha_g^2\left[1-\left(1-\beta\right)\nu_g\right]^2\sigma_g^2 \tag{A37}$$

and their true correlation coefficient is:

$$\rho=\frac{1}{\sigma_A\sigma_S}\sum_{g=1}^{G}\alpha_g^2\left[1-\left(1-\beta\right)\nu_g\right]\left(1-\nu_g\right)\sigma_g^2. \tag{A38}$$

Voters correctly perceive the variance of the signal σ_S^2, but they misperceive the variance of ability and thus both the variance of its welfare value and its correlation with the signal:

$$\tilde{\sigma}_A^2=\sum_{g=1}^{G}\alpha_g^2\left[1-\left(1-\beta\right)\nu_g\right]\sigma_g^2 \text{ and } \tilde{\rho}=\frac{\sigma_S}{\tilde{\sigma}_A}. \tag{A39}$$

Up to the proportionality constant $\phi\left(0\right)=1/\sqrt{2\pi}$, voters' mistake is thus:

$$\Delta\propto\sqrt{\sum_{g=1}^{G}\alpha_g^2\left(1-\nu_g\right)\sigma_g^2}-\sqrt{\sum_{g=1}^{G}\alpha_g^2\left(1-\nu_g+\beta\nu_g\right)\sigma_g^2}+\beta\frac{\sum_{g=1}^{G}\alpha_g^2\left(1-\nu_g+\beta\nu_g\right)\nu_g\sigma_g^2}{\sqrt{\sum_{g=1}^{G}\alpha_g^2\left(1-\nu_g+\beta\nu_g\right)^2\sigma_g^2}}. \tag{A40}$$

Up to another proportionality constant $\sqrt{\sum_{g=1}^{G}\alpha_g^2\sigma_g^2}$, it can be rewritten:

$$\Delta\propto\sqrt{1-\bar{\nu}}-\sqrt{1-\bar{\nu}+\beta\bar{\nu}}+\frac{\beta\bar{\nu}\left[1-\bar{\nu}+\beta\bar{\nu}-\left(1-\beta\right)\left(1-\bar{\nu}\right)\zeta\right]}{\sqrt{\left(1-\bar{\nu}+\beta\bar{\nu}\right)^2+\left(1-\beta\right)^2\bar{\nu}\left(1-\bar{\nu}\right)\zeta}}, \tag{A41}$$

such that $\partial\Delta/\partial\zeta<0$.

We can rewrite

$$\Delta\propto\sqrt{1-\bar{\nu}}-\sqrt{1-\bar{\nu}+\beta\bar{\nu}}+\frac{\beta\bar{\nu}}{\sqrt{1-\bar{\nu}}}Z\left(\beta,\bar{\nu},\zeta\right), \tag{A42}$$

for an auxiliary function

$$Z\left(\beta,\bar{\nu},\zeta\right)\equiv\frac{\left[\beta+\left(1-\beta\right)\left(1-\bar{\nu}\right)\left(1-\zeta\right)\right]\sqrt{1-\bar{\nu}}}{\sqrt{\left(1-\bar{\nu}+\beta\bar{\nu}\right)^2+\left(1-\beta\right)^2\bar{\nu}\left(1-\bar{\nu}\right)\zeta}}\in\left[0,1\right], \tag{A43}$$

with unambiguous derivatives:

$$\frac{\partial Z}{\partial \beta} = \frac{\zeta Z^3}{\left[\beta + (1-\beta)(1-\bar{\nu})(1-\zeta)\right]^3} > 0, \tag{A44}$$

such that $\lim_{\beta \to 0} Z = (1-\bar{\nu})(1-\zeta)/\sqrt{1-\bar{\nu}+\bar{\nu}\zeta}$ and $\lim_{\beta \to 1} Z = \sqrt{1-\bar{\nu}}$;

$$\frac{\partial Z}{\partial \bar{\nu}} = -\frac{(1-\bar{\nu}+\beta\bar{\nu})^3(1-\zeta) + \beta^3\zeta + (1+\bar{\nu})(1-\beta)^3(1-\bar{\nu})^2\zeta(1-\zeta)}{2(1-\bar{\nu})^2\left[\beta + (1-\beta)(1-\bar{\nu})(1-\zeta)\right]^3} Z^3 < 0, \tag{A45}$$

such that $\lim_{\bar{\nu} \to 0} Z = 1 - \zeta + \beta\zeta$ and $\lim_{\bar{\nu} \to 1} Z = 0$; and

$$\frac{\partial Z}{\partial \zeta} = -\frac{1}{2}(1-\beta)\frac{(2-\bar{\nu}+\beta\bar{\nu})(1-\bar{\nu}+\beta\bar{\nu}) + (1-\beta)^2\bar{\nu}(1-\bar{\nu})\zeta}{\left[\beta + (1-\beta)(1-\bar{\nu})(1-\zeta)\right]^3} Z^3 < 0 \tag{A46}$$

such that $\lim_{\zeta \to 0} Z = \sqrt{1-\bar{\nu}}$ and $\lim_{\zeta \to 1} Z = \beta\sqrt{1-\bar{\nu}}/\sqrt{1-\bar{\nu}+\beta^2\bar{\nu}}$.

Then $\Delta > 0$ if and only if

$$2Z + \frac{\beta\bar{\nu}}{1-\bar{\nu}} Z^2 > 1, \tag{A47}$$

namely if and only if $\beta > B(\bar{\nu}, \zeta)$. By the implicit-function theorem:

$$\frac{\partial B}{\partial \zeta} = -\frac{2(1-\bar{\nu}+B\bar{\nu}Z)\frac{\partial Z}{\partial \zeta}}{\bar{\nu}Z^2 + 2(1-\bar{\nu}+B\bar{\nu}Z)\frac{\partial Z}{\partial \beta}} > 0 \tag{A48}$$

and

$$\frac{\partial B}{\partial \bar{\nu}} = -\frac{BZ^2 + 2(1-\bar{\nu})(1-\bar{\nu}+B\bar{\nu}Z)\frac{\partial Z}{\partial \bar{\nu}}}{(1-\bar{\nu})\left[\bar{\nu}Z^2 + 2(1-\bar{\nu}+B\bar{\nu}Z)\frac{\partial Z}{\partial \beta}\right]} > 0. \tag{A49}$$

To establish the sign of $\partial B/\partial \bar{\nu}$, note that the definition of $\partial Z/\partial \bar{\nu}$ implies that $\partial B/\partial \bar{\nu} > 0$ if and only if:

$$\frac{Z}{B} + \frac{\bar{\nu}}{1-\bar{\nu}} Z^2 > 1 - \frac{\left[2 + B\zeta + (1-\bar{\nu}+B\bar{\nu})(1-\zeta)\right](1-B)^2(1-\bar{\nu})^2\zeta(1-\zeta)}{B^3\zeta + (1-\bar{\nu}+B\bar{\nu})^3(1-\zeta) + (1-B)^3(1+\bar{\nu})(1-\bar{\nu})^2\zeta(1-\zeta)}. \tag{A50}$$

The left-hand side is monotone decreasing in ζ and remains greater than one for $\zeta \to 1$.

In the limit, $\lim_{\beta \to 0} \Delta > 0$ if and only if

$$\frac{(1-\bar{\nu})(1-\zeta)}{\sqrt{1-\bar{\nu}+\bar{\nu}\zeta}} > \frac{1}{2}. \tag{A51}$$

A.9. Proof of Lemma 1

Suppose that in their first term Honest politicians play a mixed strategy with positive probabilities of extracting r' and $r'' > r'$. Then they must be indifferent between the two levels of rent:

$$p(r'') - p(r') = -\frac{\check{\gamma}}{\upsilon}(r'' - r') > 0 \Leftrightarrow \pi(r'') > \pi(r'). \tag{A52}$$

This implies that both Corrupt politicians and Opportunists strictly prefer r'' to r', so only Honest politicians extract rent r' with positive probability. Then $\pi(r') = 1 \geq \pi(r'')$, a contradiction. Thus, Honest politicians must play a pure strategy

Suppose that Honest politicians extract rent $r_H > 0$ in their first term. Then they must weakly prefer this positive level of rent extraction to any lower level of rent extraction:

$$p(r_H) - p(r') \geq -\frac{\check{\gamma}}{\upsilon}(r_H - r') > 0 \text{ for all } r' < r_H. \tag{A53}$$

This implies that both Corrupt politicians and Opportunists strictly prefer r_H to any $r' < r_H$, so r_H must be the minimum level of rent extraction on the equilibrium path.

The re-election probability that makes Honest politicians indifferent between their equilibrium payoff and the payoff from no rent extraction is

$$\bar{p}_H(0) = p(r_H) + \frac{\check{\gamma}}{\upsilon} r_H. \tag{A54}$$

Non-Honest politicians must have an equilibrium payoff that is at least as high as they would obtain by playing r_H. Thus, the re-election probability that makes them indifferent between their equilibrium payoff and the payoff from no rent extraction is

$$\bar{p}_i(0) \geq p(r_H) + \frac{\gamma_{i,1}}{\upsilon + \hat{r}\mathbb{E}\gamma_{i,2}} r_H > \bar{p}_H(0). \tag{A55}$$

As a consequence, our refinement ensures that voters infer $\pi(0) = 1$, which implies $p(0) \geq p(r_H)$, a contradiction. Thus, Honest politicians must play the pure strategy $r_1 = 0$ in their first term.

A.10. Proof of Lemma 2 and Proposition 6

Plugging in Equations (20), an equilibrium is defined by a system of two equations in two unknowns γ^* and η^*:

$$\gamma^* = \delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\left[F(\eta^*) - F\left(\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right)\right] \tag{A56}$$

and

$$\begin{aligned}\eta^* = {} & [1 - \pi_0 - \kappa H(\gamma^*)]\,\delta \int_{\eta^*}^{\infty} (\eta - \eta^*)\, dF(\eta) \\ & + [\pi_0 + \kappa H(\gamma^*)]\left\{\hat{r} + \delta \int_{\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}}^{\infty} \left[\eta - \eta^* + \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right] dF(\eta)\right\}.\end{aligned} \tag{A57}$$

Equation (A56) implicitly defines $\gamma^* \geq 0$ as a function of η^* with derivative:

$$\frac{\partial \gamma^*}{\partial \eta^*} = \frac{\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\left[f(\eta^*) - f\left(\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right)\right]}{1 + \delta(\upsilon + \bar{\gamma}\hat{r})\, f\left(\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right)\frac{\pi_0 \kappa h(\gamma^*)}{[\pi_0 + \kappa H(\gamma^*)]^2}}. \tag{A58}$$

Taking into account the uniform distribution of ability, this function is constant at zero for $\eta^* \leq -\hat{\eta}$ and for $\eta^* \geq \hat{\eta} + \hat{r}$. There is a value $\eta_0 \in [-\hat{\eta}, \hat{r} - \hat{\eta}]$, implicitly defined by

$$\eta_0 - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H\left(\frac{1}{2}\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\left(1 + \frac{\eta_0}{\hat{\eta}}\right)\right)} = -\hat{\eta}, \tag{A59}$$

such that for $-\hat{\eta} < \eta^* < \min\{\hat{\eta}, \eta_0\}$ the function is linear and strictly increasing:

$$\gamma^* = \frac{1}{2}\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\left(1 + \frac{\eta^*}{\hat{\eta}}\right); \tag{A60}$$

while for $\max\{\hat{\eta}, \eta_0\} < \eta^* < \hat{\eta} + \hat{r}$ the function is strictly decreasing and implicitly defined by

$$\gamma^* = \frac{1}{2}\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\left[1 - \frac{\eta^*}{\hat{\eta}} + \frac{\pi_0}{\pi_0 + \kappa H(\gamma^*)}\frac{\hat{r}}{\hat{\eta}}\right]. \tag{A61}$$

If $\eta_0 < \hat{\eta}$, then for $\eta_0 \leq \eta^* \leq \hat{\eta}$ the function is constant at γ_0^*, implicitly defined by

$$\gamma_0^*\left[\pi_0 + \kappa H(\gamma_0^*)\right] = \frac{1}{2}\delta\frac{\upsilon + \bar{\gamma}\hat{r}}{\hat{\eta}}\pi_0; \tag{A62}$$

if instead $\eta_0 > \hat{\eta}$, then for $\hat{\eta} \leq \eta^* \leq \eta_0$ the function is constant at $\delta(\upsilon/\hat{r} + \bar{\gamma})$.

Equation (A57) implicitly defines $\eta^* > 0$ as a monotone weakly increasing function of γ^*:

$$\frac{\partial \eta^*}{\partial \gamma^*} = \frac{\kappa h(\gamma^*)\left[\hat{r} + \delta\int_{\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}}^{\eta^*}(\eta - \eta^*)\, dF(\eta)\right]}{1 + \delta\left\{1 - F(\eta^*) + \left[\pi_0 + \kappa H(\gamma^*)\right]\left[F(\eta^*) - F\left(\eta^* - \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right)\right]\right\}} \geq 0, \tag{A63}$$

which is strictly increasing if and only if $0 < \gamma^* < \hat{\gamma}$. Both its minimum and its maximum can be either smaller or greater than $\hat{\eta}$.

If (but not only if) $\eta_0 \leq 0$, namely

$$\frac{\hat{\eta}}{\hat{r}} \geq \frac{\pi_0}{\pi_0 + \kappa H\left(\frac{1}{2}\delta\left(\frac{\upsilon}{\hat{r}} + \bar{\gamma}\right)\right)}, \tag{A64}$$

then Equations (A56) and (A57) define a unique equilibrium: $\gamma^* > 0$ and $\eta^* > 0$.

If $\eta_0 \leq 0$, a fortiori $\eta_0 < \hat{\eta}$. If and only if

$$\frac{\hat{\eta}}{\hat{r}} \geq \pi_0 + \kappa H(\gamma^*) + \frac{\delta}{4}\frac{\pi_0^2}{\pi_0 + \kappa H(\gamma^*)}\frac{\hat{r}}{\hat{\eta}}, \tag{A65}$$

then the unique equilibrium is $\gamma^* = \gamma_0^*$ and $\eta^* \leq \hat{\eta}$ such that

$$\begin{aligned}\eta^* = {} & \left[1 - \pi_0 - \kappa H(\gamma^*)\right]\frac{\delta}{4\hat{\eta}}(\hat{\eta} - \eta^*)^2 \\ & + \left[\pi_0 + \kappa H(\gamma^*)\right]\left\{\hat{r} + \frac{\delta}{4\hat{\eta}}\left[\hat{\eta} - \eta^* + \frac{\pi_0 \hat{r}}{\pi_0 + \kappa H(\gamma^*)}\right]^2\right\}.\end{aligned} \tag{A66}$$

Equation (A62) implies that $\partial\gamma_0^*/\partial\left(\hat{\eta}/\hat{r}\right) < 0$. Thus, Equation (A65) defines a minimum threshold for $\hat{\eta}/\hat{r}$ because its right-hand side has derivative:

$$-\frac{\delta}{4}\frac{\pi_0^2}{\pi_0+\kappa H\left(\gamma^*\right)}\left(\frac{\hat{r}}{\hat{\eta}}\right)^2-\left\{1-\frac{\delta}{4}\left[\frac{\pi_0}{\pi_0+\kappa H\left(\gamma^*\right)}\right]^2\frac{\hat{r}}{\hat{\eta}}\right\}\kappa h\left(\gamma_0^*\right)\frac{\partial\gamma_0^*}{\partial\left(\hat{\eta}/\hat{r}\right)}\leq 0, \tag{A67}$$

where the term in curly brackets is unambiguously positive because Lemma 2 ensures that $\eta^*-\pi_0\hat{r}/\left[\pi_0+\kappa H\left(\gamma^*\right)\right]\geq-\hat{\eta}$ while Equation (A65) ensures that $\eta^*\leq\hat{\eta}$. The threshold Ξ in Proposition 6 is the higher between this threshold and the right-hand side of Equation (2).

A.11. Proof of Propositions 7 and 8

The comparative statics for γ^* are immediate from Equation (A62), except:

$$\frac{\partial\ln\gamma^*}{\partial\beta}=\frac{\kappa H\left(\gamma^*\right)}{\pi_0+\kappa H\left(\gamma^*\right)+\kappa\gamma^*h\left(\gamma^*\right)}\left(\frac{\partial\ln\pi_0}{\partial\beta}-\frac{\partial\ln\kappa}{\partial\beta}\right). \tag{A68}$$

A.12. Proof of Proposition 9

We can rewrite Equation (A66) as the implicit definition of p_r:

$$1-2p_r=\left[1-\pi_0-\kappa H\left(\gamma^*\right)\right]\delta p_r^2+\left[\pi_0+\kappa H\left(\gamma^*\right)\right]\left\{\frac{\hat{r}}{\hat{\eta}}+\delta\left[p_r+\frac{1}{2}\frac{\pi_0}{\pi_0+\kappa H\left(\gamma^*\right)}\frac{\hat{r}}{\hat{\eta}}\right]^2\right\}, \tag{A69}$$

such that

$$\frac{\partial p_r}{\partial\gamma^*}=-\frac{1}{2}\frac{\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)-\delta p_0\pi_1\frac{\hat{r}}{\hat{\eta}}}{1+\left(1-\frac{\pi_0}{\pi_1}\right)\delta p_r+\frac{\pi_0}{\pi_1}\delta p_0}\kappa h\left(\gamma^*\right)<0. \tag{A70}$$

It follows immediately that $\partial p_r/\partial\delta<0$, $\partial p_r/\partial\upsilon<0$, $\partial p_r/\partial\hat{r}<0$ and $\partial p_r/\partial\hat{\eta}>0$, since the direct effect of these parameters is reinforced by their indirect effect through changes in γ^*.

A.13. Proof of Proposition 10

As voters' priors shift:

$$\begin{aligned}\frac{\partial p_r}{\partial\beta}&=-\frac{1}{2}\frac{\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]\frac{\pi_0}{\pi_1}\left(\frac{\partial\ln\pi_0}{\partial\beta}-\frac{\partial\ln\pi_1}{\partial\beta}\right)+\frac{\pi_0}{\pi_1}\delta p_0\frac{\hat{r}}{\hat{\eta}}\frac{\partial\pi_1}{\partial\beta}}{1+\left(1-\frac{\pi_0}{\pi_1}\right)\delta p_r+\frac{\pi_0}{\pi_1}\delta p_0}\\&=-\frac{1}{2}\frac{\pi_0}{\pi_1}\frac{\begin{array}{r}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]\left\{\left[1+\frac{\kappa\gamma^*h\left(\gamma^*\right)}{\pi_0}\right]\frac{\partial\pi_0}{\partial\beta}+H\left(\gamma^*\right)\frac{\partial\kappa}{\partial\beta}\right\}\\+\delta p_0\frac{\hat{r}}{\hat{\eta}}\left[\left(1-\pi_1\right)\frac{\partial\pi_0}{\partial\beta}-\pi_1H\left(\gamma^*\right)\frac{\partial\kappa}{\partial\beta}\right]\end{array}}{\left[1+\left(1-\frac{\pi_0}{\pi_1}\right)\delta p_r+\frac{\pi_0}{\pi_1}\delta p_0\right]\left[\pi_0+\kappa H\left(\gamma^*\right)+\kappa\gamma^*h\left(\gamma^*\right)\right]},\end{aligned} \tag{A71}$$

recalling that in equilibrium $\partial \ln \pi_1/\partial\beta = \partial \ln \gamma^*/\partial\beta$. Hence, $\partial p_r/\partial\beta < 0$ if and only if $\partial\kappa/\partial\beta + \Psi_r \partial\pi_0/\partial\beta > 0$, for

$$\Psi_r \equiv \frac{\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right]\left[1 + \frac{\kappa\gamma^* h(\gamma^*)}{\pi_0}\right] + \delta p_0 \frac{\hat{r}}{\hat{\eta}}\left(1 - \pi_1\right)}{\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right) - \delta p_0 \frac{\hat{r}}{\hat{\eta}} \pi_1\right] H\left(\gamma^*\right)} > 1. \tag{A72}$$

A.14. Proof of Proposition 11

We can rewrite Equation (A66) as the implicit definition of p_0:

$$\begin{aligned} 1 - 2p_0 + \frac{\pi_0}{\pi_0 + \kappa H\left(\gamma^*\right)}\frac{\hat{r}}{\hat{\eta}} = \left[1 - \pi_0 - \kappa H\left(\gamma^*\right)\right]\delta\left[p_0 - \frac{1}{2}\frac{\pi_0}{\pi_0 + \kappa H\left(\gamma^*\right)}\frac{\hat{r}}{\hat{\eta}}\right]^2 \\ + \left[\pi_0 + \kappa H\left(\gamma^*\right)\right]\left(\frac{\hat{r}}{\hat{\eta}} + \delta p_0^2\right), \end{aligned} \tag{A73}$$

such that

$$\frac{\partial p_0}{\partial \gamma^*} = -\frac{1}{2}\frac{\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right) + \left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\frac{\pi_1^2}{\pi_0}\frac{\hat{r}}{\hat{\eta}}}{1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r + \frac{\pi_0}{\pi_1}\delta p_0}\kappa h\left(\gamma^*\right) < 0. \tag{A74}$$

It follows immediately that $\partial p_r/\partial\delta < 0$ and $\partial p_r/\partial\upsilon < 0$. However,

$$\frac{\partial p_0}{\partial\left(\hat{r}/\hat{\eta}\right)} = \frac{1}{2}\frac{\left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\pi_1 - \frac{\pi_0}{\pi_1}}{1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r + \frac{\pi_0}{\pi_1}\delta p_0}. \tag{A75}$$

A.15. Proof of Proposition 12

As voters' priors shift:

$$\begin{aligned} \frac{\partial p_0}{\partial\beta} &= -\frac{1}{2}\frac{\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right]\frac{\pi_0}{\pi_1}\left(\frac{\partial\ln\pi_0}{\partial\beta} - \frac{\partial\ln\pi_1}{\partial\beta}\right) - \left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\frac{\partial\pi_1}{\partial\beta}}{1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r + \frac{\pi_0}{\pi_1}\delta p_0} \\ &= -\frac{1}{2}\frac{\begin{array}{c}\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right]\frac{\pi_0}{\pi_1}\left\{\left[1 + \frac{\kappa\gamma^* h(\gamma^*)}{\pi_0}\right]\frac{\partial\pi_0}{\partial\beta} + H\left(\gamma^*\right)\frac{\partial\kappa}{\partial\beta}\right\} \\ - \left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\left[\left(1 - \pi_1\right)\frac{\partial\pi_0}{\partial\beta} - \pi_1 H\left(\gamma^*\right)\frac{\partial\kappa}{\partial\beta}\right]\end{array}}{\left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r + \frac{\pi_0}{\pi_1}\delta p_0\right]\left[\pi_0 + \kappa H\left(\gamma^*\right) + \kappa\gamma^* h\left(\gamma^*\right)\right]}. \end{aligned} \tag{A76}$$

Hence, $\partial p_0/\partial\beta < 0$ if and only if $\partial\kappa/\partial\beta + \Psi_0 \partial\pi_0/\partial\beta > 0$, for

$$\Psi_0 \equiv \frac{\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right]\frac{\pi_0}{\pi_1}\left[1 + \frac{\kappa\gamma^* h(\gamma^*)}{\pi_0}\right] - \left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\left(1 - \pi_1\right)}{\left\{\left[\frac{\hat{r}}{\hat{\eta}} + \delta\left(p_0^2 - p_r^2\right)\right]\frac{\pi_0}{\pi_1} + \left[1 + \left(1 - \frac{\pi_0}{\pi_1}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\pi_1\right\} H\left(\gamma^*\right)} < \Psi_r. \tag{A77}$$

A.16. Proof of Corollary 4

As voters' perception of opportunism disappears, $\lim_{\kappa\to 0}\pi_1 = 1$. Thus

$$\lim_{\kappa\to 0}\frac{\partial p_0}{\partial\left(\hat{r}/\hat{\eta}\right)} = \frac{1}{2}\frac{\left(1-\pi_0\right)\left(1+\delta p_r\right)}{1+\left(1-\pi_0\right)\delta p_r+\pi_0\delta p_0} > 0 \tag{A78}$$

for all $\pi_0 < 1$, while

$$\lim_{\kappa\to 0}\frac{\partial p_0}{\partial\pi_0} = -\frac{1}{2}\frac{\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]\pi_0\left[1+\frac{\kappa\gamma^* h(\gamma^*)}{\pi_0}\right]}{\left[1+\left(1-\pi_0\right)\delta p_r+\pi_0\delta p_0\right]\left[\pi_0+\kappa H\left(\gamma^*\right)+\kappa\gamma^* h\left(\gamma^*\right)\right]} < 0. \tag{A79}$$

A.17. Proof of Corollary 5

As voters' perception of honesty disappears, Proposition 6 implies that $\lim_{\pi_0\to 0}\gamma^* = 0$ and $\lim_{\pi_0\to 0}\pi_1 = 0$. A fortiori $\lim_{\pi_0\to 0}\left(\pi_0/\pi_1\right) = 0$ because $\pi_0/\pi_1 = \pi_0+\kappa H\left(\gamma^*\right)$. Equations (23) and (24) imply that $\lim_{\pi_0\to 0}p_r = \lim_{\pi_0\to 0}p_0$. Equations (23) and (26) imply that:

$$\lim_{\pi_0\to 0}p_r = \frac{\sqrt{1+\delta}-1}{\delta}. \tag{A80}$$

Therefore,

$$\lim_{\pi_0\to 0}\frac{\partial p_0}{\partial\left(\hat{r}/\hat{\eta}\right)} = \frac{1}{2}\lim_{\pi_0\to 0}\left(\pi_1-\frac{1}{\sqrt{1+\delta}}\frac{\pi_0}{\pi_1}\right), \tag{A81}$$

while for a uniform distribution $H\left(\gamma\right) = \gamma/\hat{\gamma}$ for $\gamma\in[0,\hat{\gamma}]$,

$$\lim_{\pi_0\to 0}\frac{\partial p_0}{\partial\pi_0} = \frac{1}{2}\frac{\hat{r}}{\hat{\eta}}\lim_{\pi_0\to 0}\frac{1-\frac{1}{\sqrt{1+\delta}}\frac{\pi_0}{\pi_1^2}}{\pi_0+2\kappa H\left(\gamma^*\right)}. \tag{A82}$$

Both derivatives are negative if and only if

$$\lim_{\pi_0\to 0}\frac{\pi_0}{\pi_1^2} > \sqrt{1+\delta}. \tag{A83}$$

By Proposition (6), with uniformly distributed greed the probability that an Opportunist refrains from first-term rent extraction is $\Gamma\equiv\gamma^*/\hat{\gamma}$, implicitly defined by:

$$\Gamma\left(\pi_0+\kappa\Gamma\right) = \delta\Upsilon\pi_0 \tag{A84}$$

for

$$\Upsilon\equiv\frac{1}{2\hat{\eta}}\left(\frac{\upsilon}{\hat{\gamma}}+\frac{\hat{r}}{2}\right). \tag{A85}$$

We can solve out:

$$\Gamma = \frac{\sqrt{\left(\pi_0+4\delta\kappa\Upsilon\right)\pi_0}-\pi_0}{2\kappa}. \tag{A86}$$

By Equation (20):

$$\frac{\pi_0}{\pi_1^2} = \pi_0\left(\frac{\delta\Upsilon}{\Gamma}\right)^2 = \frac{2\left(\delta\kappa\Upsilon\right)^2}{2\delta\kappa\Upsilon+\pi_0-\sqrt{\left(\pi_0+4\delta\kappa\Upsilon\right)\pi_0}}, \tag{A87}$$

such that

$$\lim_{\pi_0\to 0}\frac{\pi_0}{\pi_1^2} = \delta\kappa\Upsilon. \tag{A88}$$

A.18. Proof of Proposition 13

Subtracting Equation (32) from Equation (26), voters' misperception of their own welfare is:

$$\frac{\eta^*}{\hat{\eta}}-\frac{\eta^u}{\hat{\eta}}=\frac{\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]\left(\frac{\pi_0}{\pi_1}-\frac{\pi_0^u}{\pi_1^u}\right)+\frac{\pi_0^u}{\pi_1^u}\delta p_0\left(\pi_1-\pi_1^u\right)\frac{\hat{r}}{\hat{\eta}}}{1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r+\frac{\pi_0^u}{\pi_1^u}\delta p_0}$$

$$=\frac{\begin{array}{c}\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]+\frac{\pi_0^u}{\pi_1^u}\delta p_0\frac{\hat{r}}{\hat{\eta}}\left(1-\pi_1^u\right)\right\}\left(\pi_0-\pi_0^u\right)\\+\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]-\frac{\pi_0^u}{\pi_1^u}\delta p_0\frac{\hat{r}}{\hat{\eta}}\pi_1^u\right\}H\left(\gamma^*\right)\left(\kappa-\kappa^u\right)\end{array}}{\frac{\pi_0}{\pi_1}\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r+\frac{\pi_0^u}{\pi_1^u}\delta p_0\right]}. \quad \text{(A89)}$$

Voters' screening of politicians who did not extract rents differs from the optimal one-time deviation by:

$$\frac{\eta^*}{\hat{\eta}}-\frac{\eta^u}{\hat{\eta}}-\left(\pi_1-\pi_1^u\right)\frac{\hat{r}}{\hat{\eta}}=\frac{\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]\left(\frac{\pi_0}{\pi_1}-\frac{\pi_0^u}{\pi_1^u}\right)-\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\left(\pi_1-\pi_1^u\right)\frac{\hat{r}}{\hat{\eta}}}{1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r+\frac{\pi_0^u}{\pi_1^u}\delta p_0}$$

$$=\frac{\begin{array}{c}\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]-\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\left(1-\pi_1^u\right)\right\}\left(\pi_0-\pi_0^u\right)\\+\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]+\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\pi_1^u\right\}H\left(\gamma^*\right)\left(\kappa-\kappa^u\right)\end{array}}{\frac{\pi_0}{\pi_1}\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r+\frac{\pi_0^u}{\pi_1^u}\delta p_0\right]}. \quad \text{(A90)}$$

Suppose that $\pi_0 \geq \pi_0^u$. Then $\eta^* > \eta^u$ if and only if $\Xi_r\left(\pi_0-\pi_0^u\right)+\kappa-\kappa^u>0$ for

$$\Xi_r \equiv \frac{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]+\frac{\pi_0^u}{\pi_1^u}\delta p_0\frac{\hat{r}}{\hat{\eta}}\left(1-\pi_1^u\right)}{\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]-\frac{\pi_0^u}{\pi_1^u}\delta p_0\frac{\hat{r}}{\hat{\eta}}\pi_1^u\right\}H\left(\gamma^*\right)}>1, \quad \text{(A91)}$$

while $\eta^*-\pi_1\hat{r}>\eta^u-\pi_1^u\hat{r}$ if and only if $\Xi_0\left(\pi_0-\pi_0^u\right)+\kappa-\kappa^u>0$ for

$$\Xi_0 \equiv \frac{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]-\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\left(1-\pi_1^u\right)}{\left\{\frac{\pi_0}{\pi_1}\left[\frac{\hat{r}}{\hat{\eta}}+\delta\left(p_0^2-p_r^2\right)\right]+\left[1+\left(1-\frac{\pi_0^u}{\pi_1^u}\right)\delta p_r\right]\frac{\hat{r}}{\hat{\eta}}\pi_1^u\right\}H\left(\gamma^*\right)}<\Xi_r. \quad \text{(A92)}$$

A.19. Proof of Lemma 3

The discussion of Lemma 3 in the body of the text establishes that voters re-elect a rent extractor, and appoint him perpetual dictator when given the opportunity, under the same condition $\eta_i > \eta^* \equiv (1-\delta) W_1$. If the incumbent refrained from rent extraction, the threshold is $\eta^*-\pi_1\hat{r}$. An Opportunist who correctly anticipates these thresholds refrain from rent extraction in the first term according to Equation (19) for

$$\delta=\delta_0\left(1-\varepsilon+\frac{\varepsilon}{1-\delta_0}\right). \quad \text{(A93)}$$

Voter's expectation of their own welfare when electing a random politician to his first term is

$$W_1 = [1 - \pi_0 - \kappa H(\gamma^*)] \delta_0 \left\{ F(\eta^*) W_1 + \int_{\eta^*}^{\infty} \left[(1-\varepsilon)(\eta + \delta_0 W_1) + \varepsilon \frac{\eta}{1-\delta_0} \right] dF(\eta) \right\}$$
$$+ [\pi_0 + \kappa H(\gamma^*)]$$
$$\times \left(\hat{r} + \delta_0 \left\{ F(\eta^* - \pi_1 \hat{r}) W_1 + \int_{\eta^* - \pi_1 \hat{r}}^{\infty} \left[(1-\varepsilon)(\eta + \pi_1 \hat{r} + \delta_0 W_1) + \varepsilon \frac{\eta + \pi_1 \hat{r}}{1-\delta_0} \right] dF(\eta) \right\} \right). \tag{A94}$$

Plugging in $W_1 = \eta^* / (1-\delta)$ yields Equation (21).

References

[1] Johnson, Norman L., Samuel Kotz and Narayanaswamy Balakrishnan. 1994. *Continuous Univariate Distributions, Volume 1,* second edition. New York, NY: Wiley.

[2] Kotz, Samuel, Narayanaswamy Balakrishnan and Norman L. Johnson. 2000. *Continuous Multivariate Distributions, Volume 1: Models and Applications,* second edition. New York, NY: Wiley.

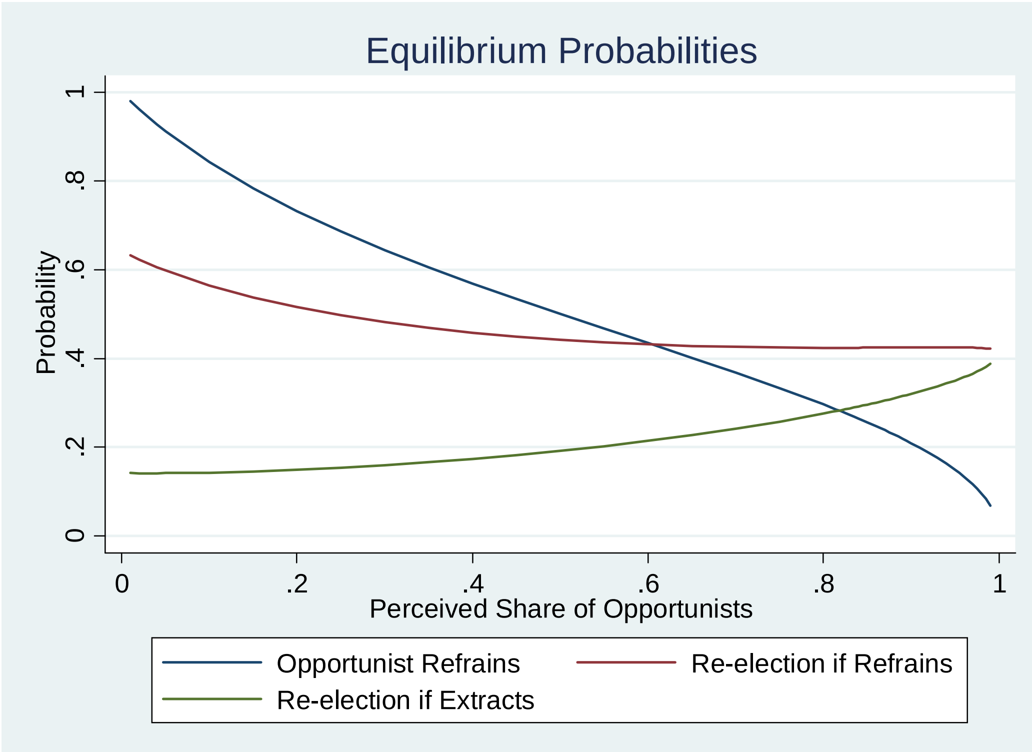

Figure 1: Equilibrium when 50% of non-Opportunists are perceived to be Honest

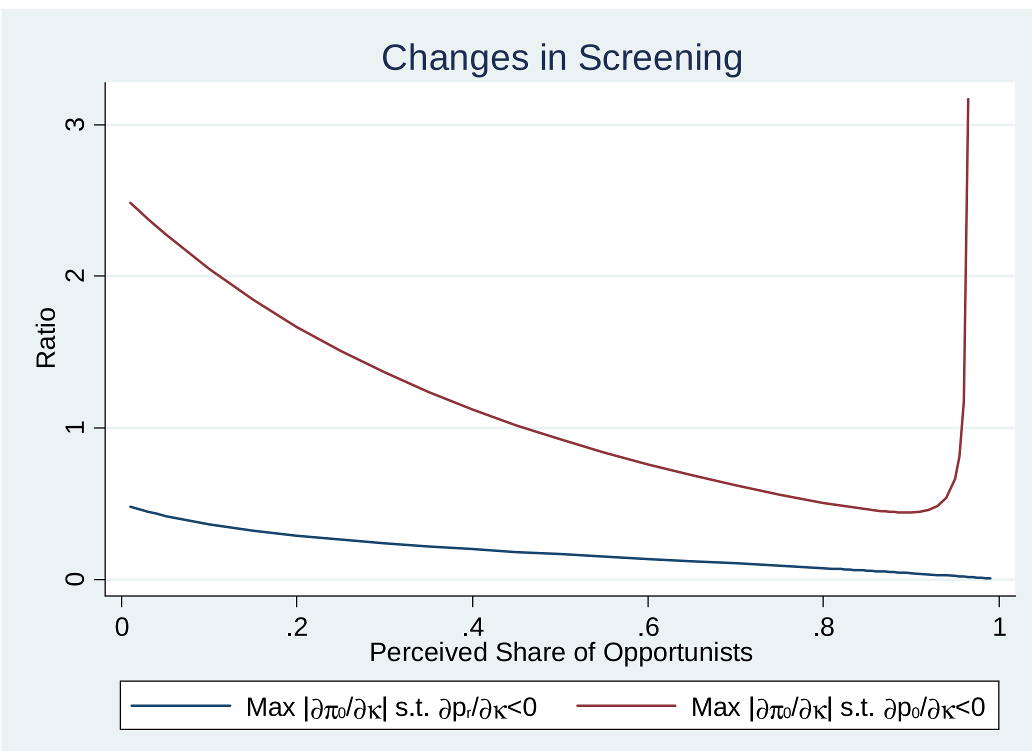

Figure 2: Selection when 50% of non-Opportunists are perceived to be Honest

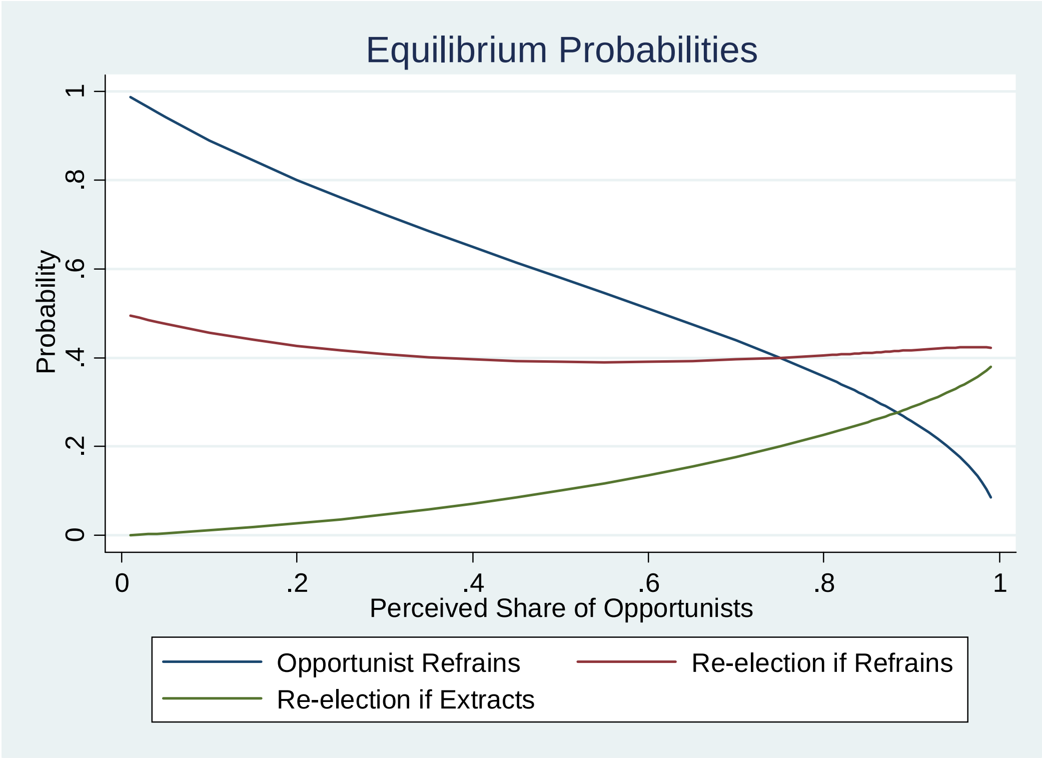

Figure 3: Equilibrium when 80% of non-Opportunists are perceived to be Honest

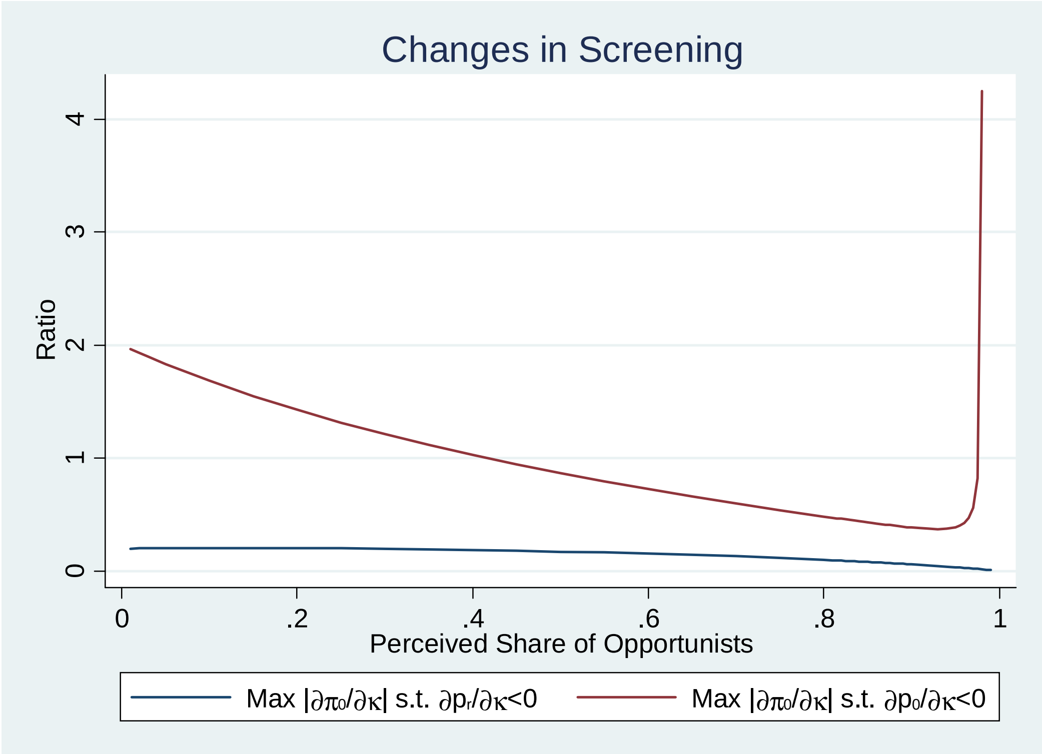

Figure 4: Selection when 80% of non-Opportunists are perceived to be Honest

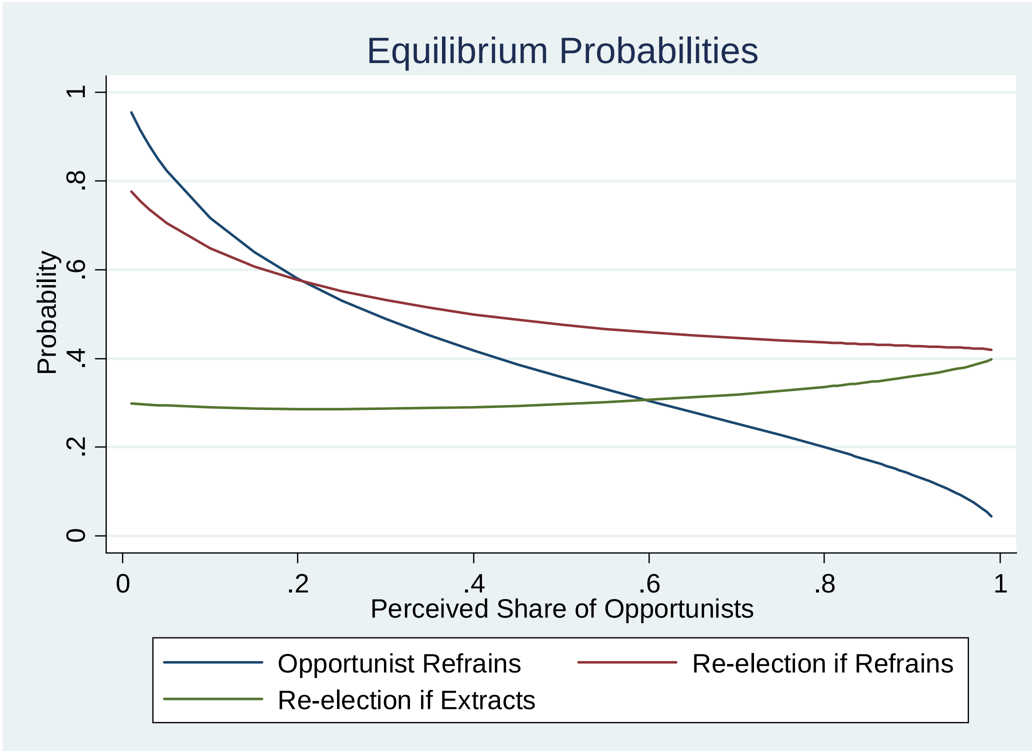

Figure 5: Equilibrium when 20% of non-Opportunists are perceived to be Honest

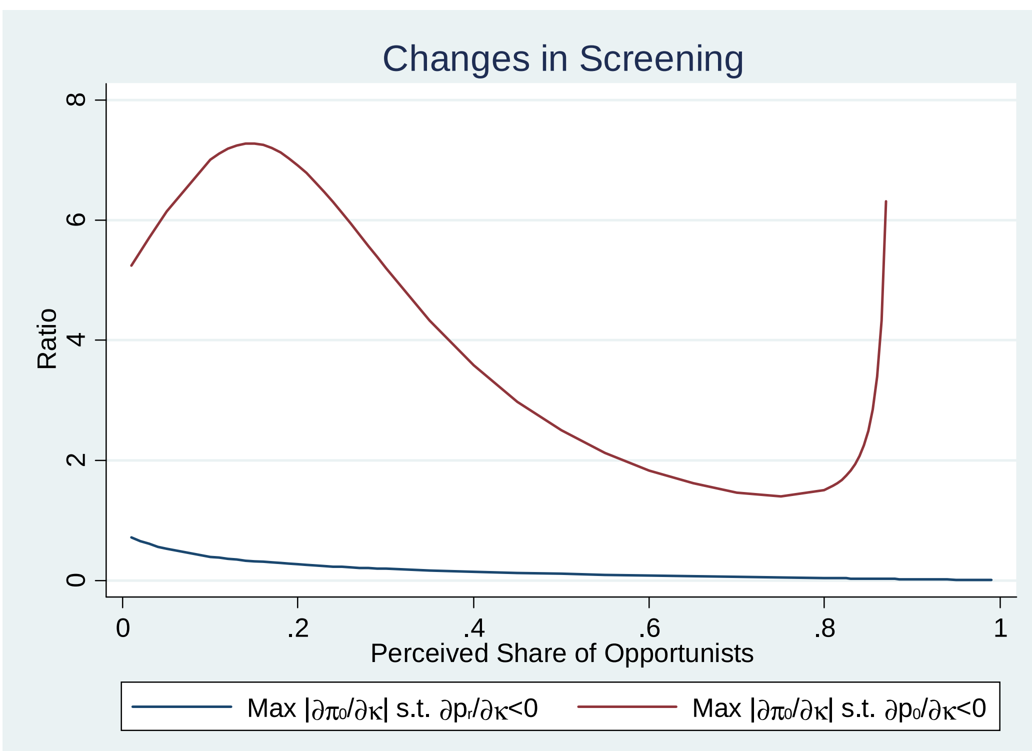

Figure 6: Selection when 20% of non-Opportunists are perceived to be Honest

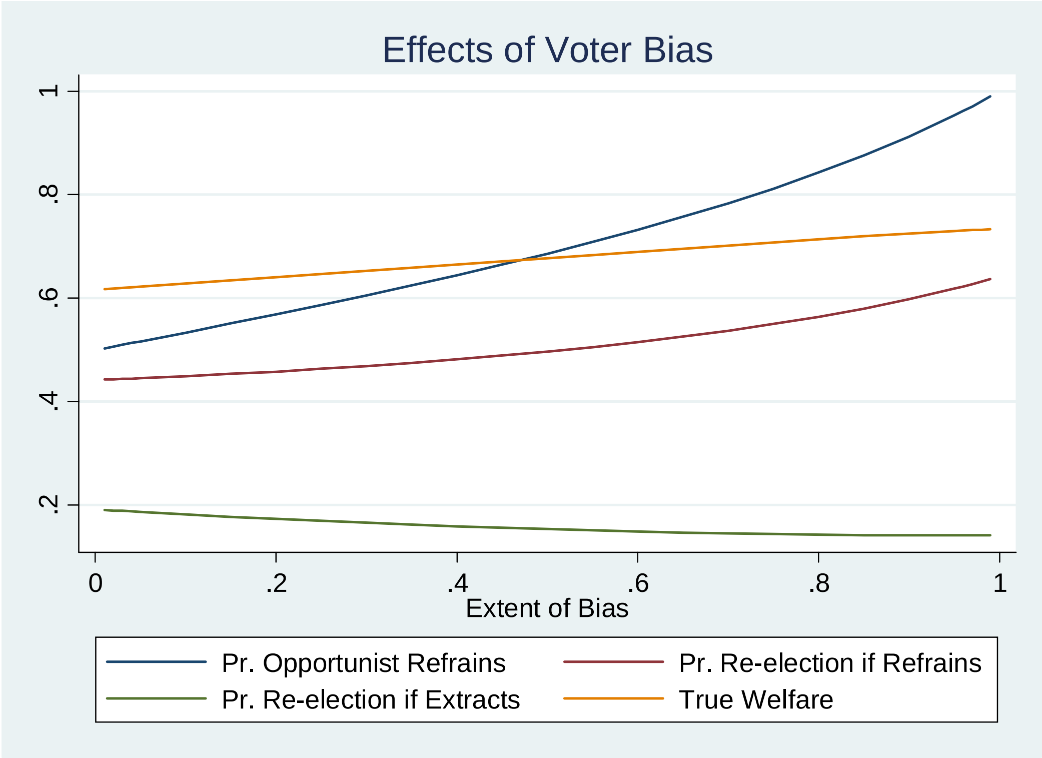

Figure 7: Welfare when 50% of politicians are Opportunists and 25% Honest

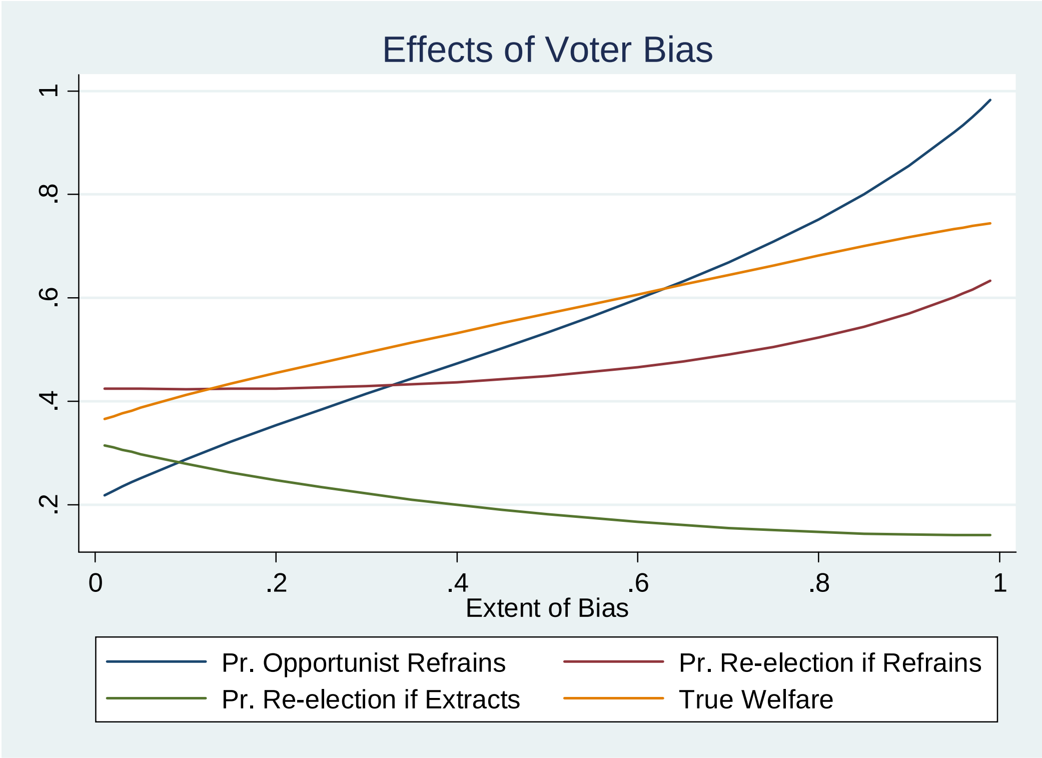

Figure 8: Welfare when 90% of politicians are Opportunists and 5% Honest

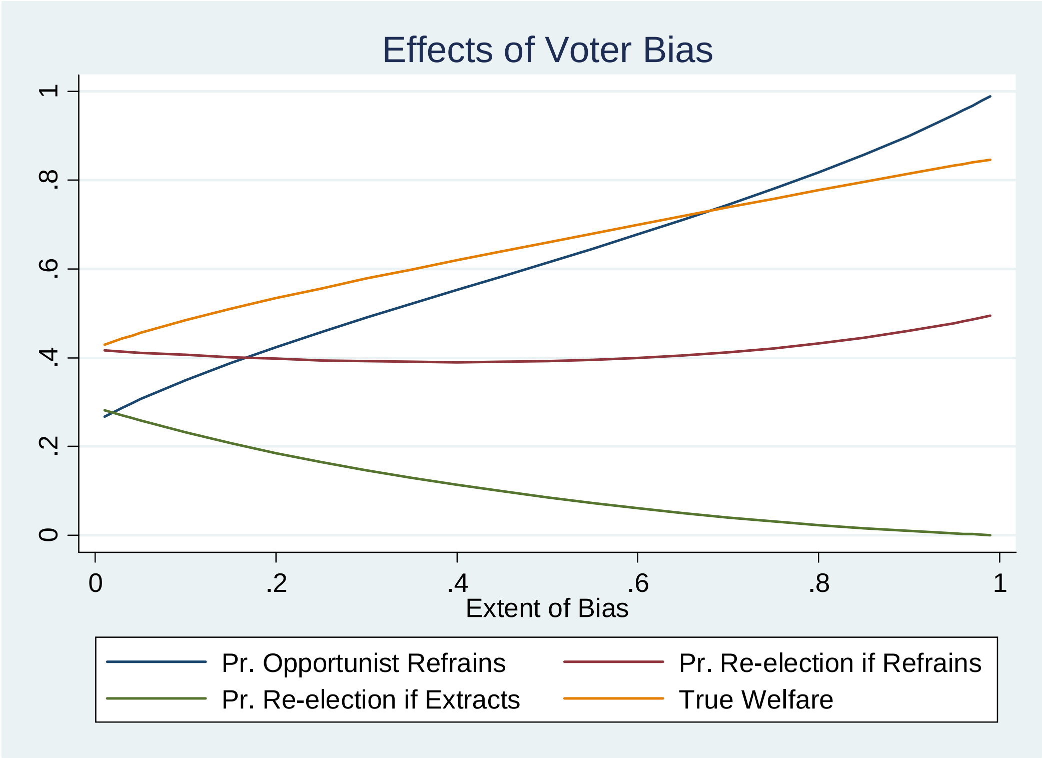

Figure 9: Welfare when 90% of politicians are Opportunists and 8% Honest

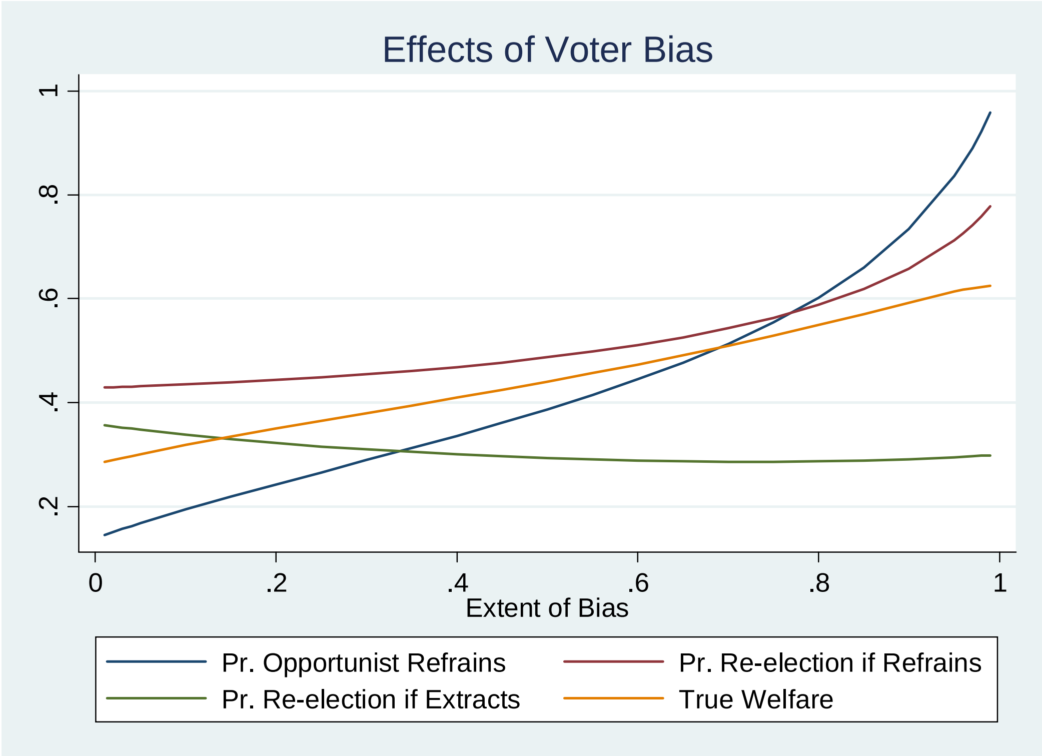

Figure 10: Welfare when 90% of politicians are Opportunists and 2% are Honest

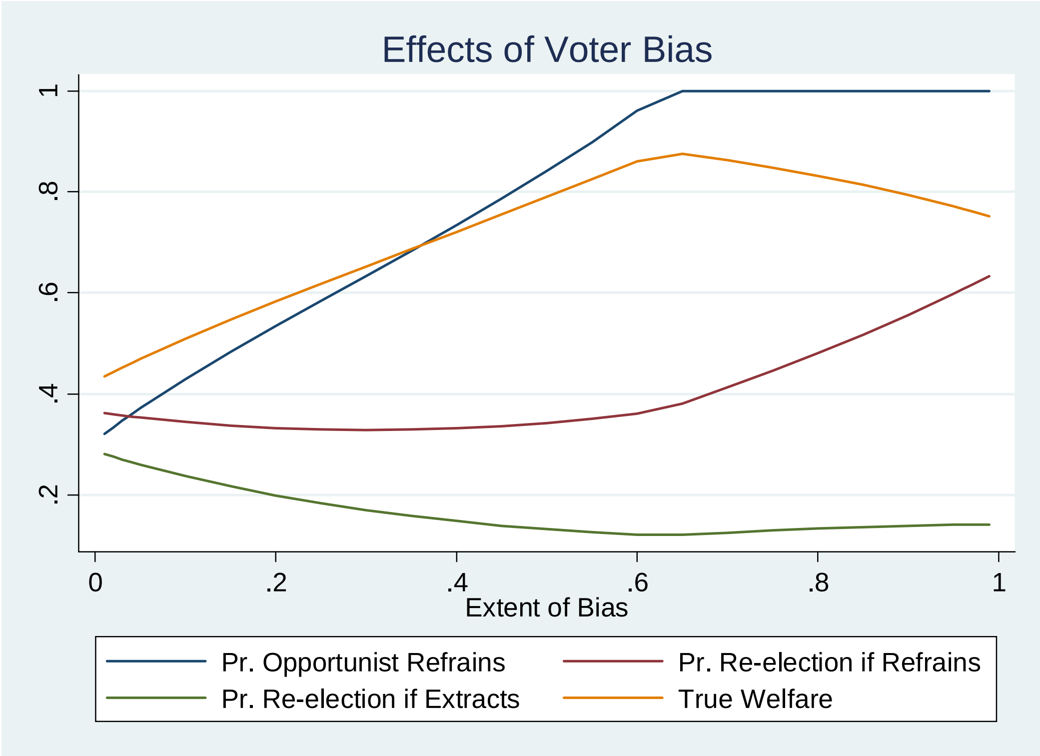

Figure 11: Welfare with a high pure value of office

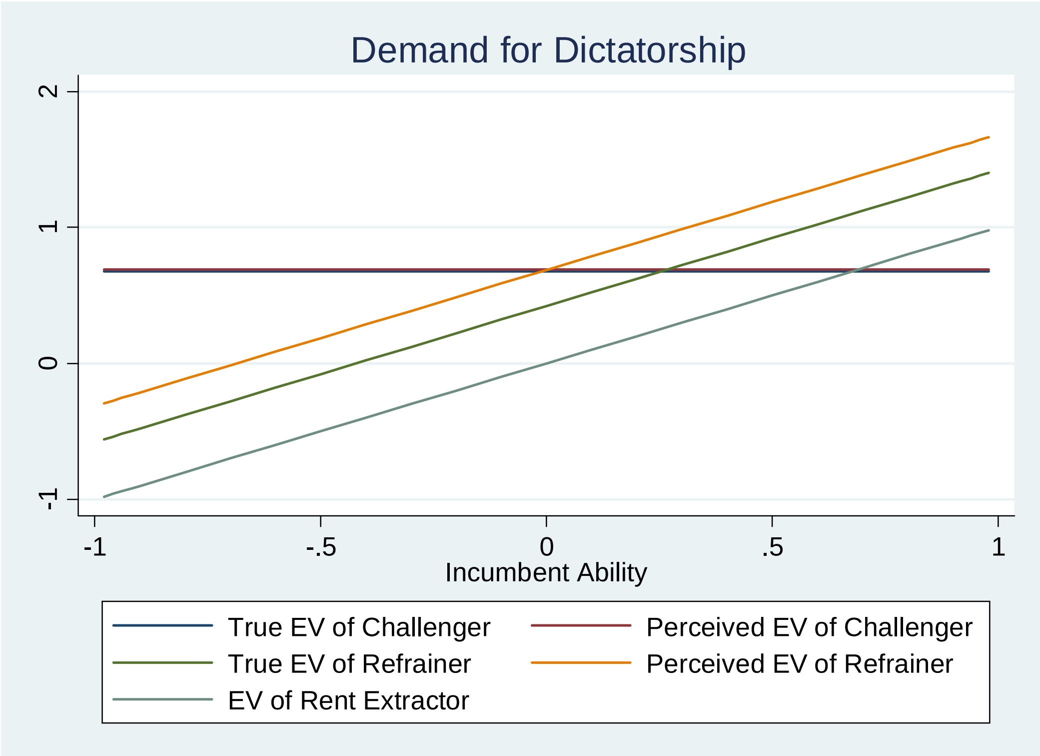

Figure 12: Comparisons when 50% of politicians are Opportunists

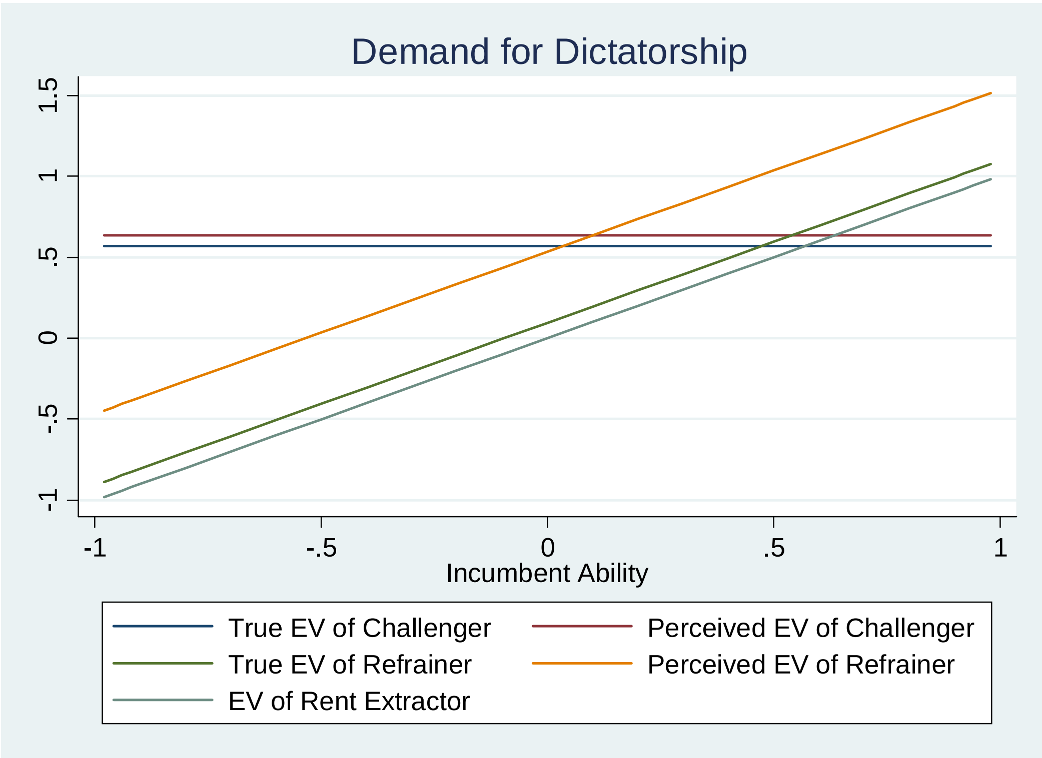

Figure 13: Comparisons when 90% of politicians are Opportunists