Revealed Political Power

Jinhui H. Bai* and **Roger Lagunoff†**

May 20, 2010‡

Abstract

This paper adopts a "revealed preference" approach to the question of what can be inferred about bias in a political system. We model an infinite horizon, dynamic economy and its political system from the point of view of an "outside observer." The observer sees a finite sequence of policy data, but does not observe the underlying distribution of political power that produced this data. Neither does he observe the preference profile of the citizenry. The observer makes inferences about distribution of political power as if political power were derived from a wealth-weighted voting system with weights that can vary across states. The weights determine the nature and magnitude of the wealth bias. Positive weights on relative income in any period indicate an "elitist" bias in the political system whereas negative weights indicate a "populist" one.

We ask: what class of weighted systems can rationalize the policy data as weighted-majority outcomes each period? We show that without further knowledge, all forms of bias are possible: *any* policy data can be shown to be rationalized by *any* system of wealth-weighted voting. An additional single crossing restriction on preferences can, however, rule out certain weighting systems. We then augment policy data with polling data and show that the set of rationalizing wealth-weights are bounded above and below, thus ruling out extreme biases. In some cases, polls can provide information about the change in political inequality across time.

JEL Codes: C73, D63, D72, D74, H11

Key Words and Phrases: wealth-bias, elitist bias, populist bias, weighted majority winner, rationalizing weights, "Anything Goes Theorems".

*Department of Economics, Georgetown University, Washington, DC 20057, USA. E-mail: jb543@georgetown.edu, Phone: +1-202-687-0935, Fax: +1-202-687-6102. USA.

†Department of Economics, Georgetown University, Washington, DC 20057, USA. 202 687-1510. lagunofr@georgetown.edu. www.georgetown.edu/lagunoff/lagunoff.htm .

‡We thank seminar participants at Indiana, Maryland, Northwestern, UC Santa Barbara, Towson, the Summer Econometric Society Meetings, the Stanford-Kingston Political Economy Conference, and the Midwest Theory Conference for helpful comments and suggestions.

1 Introduction

The principle of *political equality* is widely accepted as a governing philosophy in most democracies. According to the principle, all individuals, regardless of income or background should be endowed with the same political power or influence. On paper, electoral processes in most democracies satisfy some rough form of it, often taking the form of "one-man-one-vote" electoral systems. Examples include Winner-take-all Presidential elections (in the U.S. and Latin America) and Proportional Representation in Parliamentary elections (e.g., Western Europe).[1]

It is unlikely, however, that the *de facto* distribution of power in these countries is equal. There is anecdotal evidence, and some systematic evidence, that wealth matters in the political process. For instance, Rosenstone and Hansen (1993) show that the propensity to participate in every reported form of political activities rises with income. Campante (2008) uses campaign contribution data in the 2000 US presidential election to show that income inequality increases the share of contributions coming from relatively wealthy individuals. Bartels (2008) offers a sweeping look at the relation between economic and political inequality. He examines whether economic inequality creates political inequality in the *policy process*. Using data from the Senate Election Study, he finds that Senators' voting records are unresponsive to preferences of those in the lower third of the income distribution.[2] By contrast, Senator's responsiveness to middle and upper thirds is virtually linear to income.

These studies all suggest some form of wealth-bias in the political system. They find that the *de facto* allocation of power is such that richer individuals have a disproportionate influence in the policy process. The result is that policies enacted appear to favor wealthier rather than poorer individuals. Consequently, economic inequality apparently produces political inequality to some degree.

The present paper takes a step back by asking whether and how bias can be identified directly from policy data. When, for instance, can the egalitarian distribution of power based on "one-man-one-vote" be ruled out?

To address these issues we model an infinite horizon, dynamic economy and its political system from the point of view of an "outside observer." Policies in this economy are determined each period by a political process that aggregates the preferences of a continuum of citizens or citizen-types. The outsider observes this policy data for finite periods. He also sees or knows certain underlying attributes of the economy. He observes, for instance, the income distribution of citizen-types and, further, knows the Markov process by which it evolves in future periods. The observer does not observe, however, the preferences profile of the citizenry.

[1] Clearly, there are well known exceptions. In the U.S. representation in the Senate is equal across states, so that voters in small states have disproportionate political power in that governing body.

[2] See Chapter 9 of Bartels (2008). The Senate Election Study consists of survey data conducted after the November elections of 1988, 1990, 1992.

Instead, he knows only that long run preferences over policies are such that an individual's preferred policy is increasing in income.

The outsider's task is to infer something about the underlying distribution of political power that generated the observed policies. Specifically, he looks for sequences of distributions of political power that could have rationalized the observed policy data as majority-winning outcomes of a voting process. In this case, the voting rule allocates weights to each type's income. These weights can vary with the state of the economy and are independent of the income-generating process. Policies are chosen *as if* they came from wealth-weighted voting. In this sense, the weights correspond to an implied distribution of political power.

The weights can be either positive, indicating a pro-wealth bias, or they can be negative, indicating an anti-wealth bias. More generally, an increase in the wealth-weight works in favor of the wealthy, while a decrease works in favor of the poor. In the case of a pro-wealth bias, a wealthy individual's vote is worth more than a poorer one. We refer to this as an *elitist bias*. In the case of an anti-wealth bias, the poorer individual's vote is worth more than a richer one. We refer to this as an *populist bias*. The case where the weights are exactly zero corresponds to the standard system of "one-man-one vote" or equal representation. We refer to this as the *unbiased* system.

Given a set of weights, a "Political Lorenz Curve" can be calculated to express the implied vote share (hence, "political power") of the poorest jth portion of the population, for each possible j. A policy rule produces a *Weighted majority winner (WMW)* of the wealth-weighted voting system if in each state, the resulting policy wins in a weighted majority vote against any alternative. The system of wealth-weighted voting is then said to *rationalize the policy rule* if the policy rule produces a WMW under an admissible preference profile.

We restrict attention to policy data that are consistent with a Markov policy rule mapping states to policies. The Markov restriction focuses attention on large, anonymous polities in which reputation and other history-dependent enforcement mechanisms do not arise.[3] Given the Markov restriction, our results provide (i) necessary and sufficient conditions under which there exists a system of wealth-weights that rationalize the data, and (ii) necessary and sufficient conditions under which a *particular* weighting system rationalizes the data.

The concern in this paper with consistency of observed policy data with political fundamentals draws an obvious parallel to Revealed Preference Theory (RPT) which typically examines consistency of consumption data with budget-constrained utility maximization.[4] Our approach follows in the tradition of Afriat (1967) who examines how an individual utility function can be constructed from finite consumption and price data.[5]

[3]Further implications of the Markov restriction are discussed in Section 2.

[4]See Richter (1966) and more recently Varian (2006) for summaries and surveys of RPT developed by Paul Samuelson and others.

[5]See also Varian (1982), Chiappori and Rochet (1987), and, for a recent application of RPT to political choices, Kalandrakis (2010) .

One difference from traditional RPT is that our consistency check involves aggregation of choice: we check whether observed policies could have arisen as an equilibrium of a (possibly biased) voting process. In this sense, a closer comparison is the classic Sonnenschein-Mantel-Debreu result checking whether an aggregate excess demand function is consistent with economy-wide aggregation of optimizing choices.[6] Degan and Merlo (2009) apply consistency-of-aggregation ideas to politics. Using micro-level voting data, they examine whether the outcomes of simultaneous multi-candidate elections can be rationalized by ideological voting behavior. The consistency check in our paper bears some resemblance to their theoretical approach, except that both preferences and attributes of the political system are latent.

A second difference with most RPT models is that the policy outcomes studied here consist of a time series produced by the same underlying polity. For this reason, the present model is dynamic, putting it closer to Boldrin and Montrucchio (1986) who examine whether a given policy rule could have been rationalized by a single dynamically-consistent decision maker in a capital accumulation model.

Our first result is, in fact, reminiscent of "Anything Goes Theorems" of both Sonnenschein-Mantel-Debreu and Boldrin-Montrucchio.[7] We show that *any* policy data can be rationalized by *any* wealth bias. That is, without further structure on admissible preferences or additional forms of data, the policy data alone is not very discerning; it is consistent with every type of income-weighted bias.

This result is somewhat troubling in light of the empirical findings of Bartels and others suggesting a distinctly elitist (pro-wealth) bias. Most of theoretical literature also supports this case. One prominent theory links bias to differential participation rates among the rich and poor. Examples include Benabou (2000) and Bourguignon and Verdier (2000). In these models, the poor vote less frequently, the effect being that wealthier voters have a disproportionate influence on policy. A second type theory concerns the effect of campaign contributions, for instance, Austen-Smith (1987), Grossman and Helpman (1996), Prat (2002), Coate (2004), Campante (2008), etc. In these models, the money either "buys" influence directly or it affects policy indirectly by changing the electoral odds toward candidates ideologically predisposed toward the rich. Because contributions skew toward the wealthy, policies are biased in their favor. Finally, a third type of theory centers on disenfranchising investments, e.g., Acemoglu and Robinson (2008), made by a wealthy elite in order to disinherit the poor from the political process.

The present, "detail free" approach does not take a stand on which, if any, of these theories is the "right one." Nor do we presume that the bias should even be elitist (pro-wealth). Our first result suggests that it would be difficult to rule out any particular bias without further

[6] References for this result are Sonnenschein (1973), Mantel (1974), Debreu (1974). See also references and recent results in Brown and Kubler (2008) for applications of RPT to general equilibrium theory.

[7] The S-M-D results show that fairly weak conditions are required for consistency between aggregate demand and individual optimization. Boldrin and Montrucchio show that any capital accumulation rule is consistent with an individual's dynamic optimization.

structure on the model.

Hence, in order to develop a more meaningful inference, two further routes are taken. First, we narrow the class of admissible preferences. By imposing an additional single crossing restriction, we can rule out certain bias weights, by showing that they could not have rationalized the data. The unbiased weighting system, for instance, cannot rationalize any policy data that decreases in the state.

Second, we examine the inference problem when the observer has access to polling data. Polls provide data on specific aggregate binary orderings between benchmark policies — typically those that are being considered in the political process. Initially, we consider two simple polls in each state. In one, the observed policy is pitted against a policy located to its right (the "right-wing alternative"). In the other, the observed policy is pitted against a "left-wing alternative". The analysis is later generalized to allow for an arbitrary number of polls.

We characterize both sufficient and necessary conditions for a system of wealth weights to rationalize both the policy and poll data. It turns out that fairly minimal amounts of polling data can provide clear restrictions on the bias. Upper and lower bounds on the bias are characterized state-by-state. An upper bound represents a maximal degree of positive wealth-bias — the largest possible bias in favor of wealthy individuals. The lower bound represents the lowest possible bias.

Both bounds are shown to be computed explicitly from polling support rates. The intuition is roughly the following. Suppose a poll reveals that some fraction p_t of the population preferred the observed policy over the "right-wing" alternative at date t. Under single crossing, this implies that, say, the richest $1 - p_t$ portion of the population had a weighted vote share at date t smaller than 50%. If it were otherwise, then this richest group would have had the clout to veto the observed policy, contradicting the fact that the observed policy is a weighted majority winner. Consequently, the income weights can be no greater than that necessary to lift the $1 - p_t$ wealthiest individuals to the 50% weighted voting threshold. This gives upper bound for the bias weight, with similar intuition used to derive the lower bound.

This logic implies explicit limits for how extreme the bias can be. We show how the resulting bias bounds can be combined with period-by-period changes in poll data to examine whether political power to the wealthy increases or decreases over time.

The paper is organized as follows. Section 2 lays out the economic side of the model. Section 3 then describes the political side: an implied voting process with latent, wealth-weights. Section 4 describes the "anything goes" result, then examines the same question when citizens' preferences come from a more restrictive set. Section 5 examines the addition of poll data, using the polling to derive both static bounds and dynamic restrictions on the bias. Later in the section, more exacting assumptions are needed to examine the link between economic and political inequality. The analysis is generalized to the case of arbitrary numbers of polls. Section 6 finally concludes with a discussion of extensions. The Appendix follows.

2 The Economic Side

This section models the "economic side" of the model from the point of view of an outside observer. The tangible attributes of the economy such as income inequality, transition rules, policies, etc., are observable. The observer does not see either the parametric preferences, nor the underlying power distribution that produced the observed policies. Both the observed and unobserved attributes are laid out in the following subsections. The political side is taken up in Section 3. Throughout the paper, all functions are assumed to be measurable functions of Euclidean spaces.

2.1 The Tangible Environment

An infinite horizon economy is populated by a continuum of $I = [0, 1]$ of *citizen-types*. A citizen-type is an index that orders individuals by income, with higher types accorded higher income. A citizen of type $i \in I$ holds income $y(i, \omega_t)$ in period t that depends on the value of an aggregate state variable ω_t. For concreteness, this state can be interpreted as an economy-wide public capital stock, such as public infrastructure. The set of possible states is given by Ω, a connected subset of $\mathbb{R}$. The function y is assumed to be continuous and increasing in i, with $y(0, \omega_t) > 0$.[8]

The monotonicity of y in i means that higher citizen types are wealthier. The assumption also implies a well defined distribution function $i = h(\tilde{y}, \omega_t)$ corresponding to the proportion of types holding income no greater than $\tilde{y}$.

Each period t this society collectively determines a policy a_t. Assume $a_t \in A$ with A a compact interval in $\mathbb{R}$. The current policy and state jointly determine next period's state according to a Markov transition technology given by $\omega_{t+1} = Q(\omega_t, a_t)$, with the initial state ω_1 exogenously given. The simplest interpretation is that ω_t is the current stock of public capital such as infrastructure and a_t is an investment that augments the stock of infrastructure. There are no shocks, and $\delta \in [0, 1)$ is the common discount factor.[9]

Putting these attributes together, the physical environment is summarized by the following list $(I, \Omega, A, Q, y, \omega_1, \delta)$. The outside observer sees/knows all these attributes which remain fixed for the rest of the analysis.

[8] From here on, the term "increasing" will be taken to mean "strictly increasing", and the term "weakly increasing" will be taken to mean "nondecreasing".

[9] The special case of $\delta = 0$ corresponds to the static interpretation common in revealed preference theory.

2.2 Policy Data and Policy Preferences

The outside observer views the publicly available data over a potentially long but finite horizon of length T. Let $\{a_t\}_{t=1}^T$ denote the observed path of policies. For simplicity, the path of states $\{\omega_t\}_{t=1}^T$ is also assumed to be viewed by the outside observer, although this assumption is not necessary in the dynamic model. Given the initial state ω_1 and the observed policy path $\{a_t\}_{t=1}^T$, the state path can be easily inferred from the technological constraint, $\omega_{t+1} = Q(\omega_t, a_t)$ for each $t = 1, \dots, T-1$. In the subsequent notation, ω_t and a_t will refer to the on-path observations at date t, while ω and a will connote a generic state and policy, resp., either on the observed path or off it.

We restrict attention to *Markov* data paths, i.e., data paths for which there exist (Markov) functions $\Psi : \Omega \to A$ satisfying

$$\Psi(\omega_t) \ = \ a_t \quad \forall \ t = 1, \dots, T$$

The Markov restriction allows for a tractable characterization of the data even as it entails some loss of generality. It seems most appropriate in large and anonymous societies where history-dependent enforcement mechanisms would be difficult to implement. The fact that Markov data are consistent with single agent optimization will prove useful for comparisons later on. From here on, all data paths will be assumed to be Markov. Any function Ψ consistent with the data will be referred to as a *Markov policy rule* or just *policy rule*.

Preferences over policies are assumed to be represented by a function, $U(i, \omega, a; \Psi)$ denoting the long run payoff to a citizen-type i of (generic) policy a in (generic) state ω when future payoffs are pinned down by a policy rule Ψ. The precise form of function U is not known to the outside observer. However, U is known to belong to a set of *admissible* payoff functions defined by:

(A1) (Single Peakedness) U is continuous in the index i, and single peaked in its ath argument.

(A2) (Single Crossing) U satisfies the single crossing property in $(a\ ; i)$.[10]

(A3) (Recursive Consistency). There exist flow payoff function u such that

$$U(i, \omega, a; \Psi) \ = \ u(\omega, y(i, \omega), a) \ + \ \delta \ U(i, Q(\omega, a), \Psi(Q(\omega, a)); \Psi).$$

The single crossing property (A2) implies that in every state, wealthier citizens always prefer larger policies than poorer citizens. Assumption (A3) is a dynamic consistency restriction that also requires that flow payoffs do not depend directly on one's type; all types have

[10] A function $f(x, y)$ will be said to satisfy the *single crossing property in* $(x; y)$ if for all $x > \hat{x}$ and $y > \hat{y}$, $f(x, \hat{y}) - f(\hat{x}, \hat{y}) \ (>) \geq \ 0$ implies $f(x, y) - f(\hat{x}, y) \ (>) \ \geq \ 0$, and satisfies *strict single crossing in* $(x; y)$ if $f(x, \hat{y}) - f(\hat{x}, \hat{y}) \ \geq \ 0$ implies $f(x, y) - f(\hat{x}, y) \ > \ 0$. The "single crossing property" as defined here may be more accurately described as "single crossing from below." But because policies have no specific interpretation, notions of "larger" and "smaller" are arbitrary. Hence, without loss of generality, we could also have assumed single crossing from above.

the same underlying preferences, and consequently heterogeneity comes exclusively from differences in income which is given by data. A payoff function U satisfying (A1)-(A3) is referred to as an *admissible preference profile*.

It's worth noting that, restrictiveness in the class of admissible profiles strengthens rather than weakens certain of our results. The reason is that the larger the set of admissible preference orderings, the easier it is to find one that "works" in the sense that a political system can produce Ψ under such preferences. The narrower the class of preferences the more difficult it is for a particular system to have generated the data. Hence, possibility results (i.e., assertions that Ψ *can* be produced by a particular system) are stronger under narrower classes of preferences, while impossibility results (i.e., assertions that Ψ *cannot* be produced) are weaker, all else equal.[11]

3 The Political Side

This section specifies a class of distributions of political power, each parameterized by a "wealth bias" term. These distribution will have similar properties to standard income Lorenz curves. Each "political Lorenz curve" in this class describes a proportion of political power held by the poorest $j\%$ of the population in each state. The interpretation is that of an implied, weighted vote. Power is measured by whether and how much of a weight would one have to give to income or wealth so that the observed policy is consistent with voting.

3.1 Elitist versus Populist Bias

Political bias will be captured by a functional parameter $\alpha(\omega)$ that measures the extent of the bias in each state. Roughly, larger values of $\alpha(\omega)$ will correspond to greater political weight accorded to the rich. The weight $\alpha(\omega)$ is a parameter in a continuous integrable function $\lambda : \mathbb{R}^3 \to \mathbb{R}_+$ with the following interpretation. The value $\lambda(\tilde{y}, \tilde{\alpha}, \omega)$ connotes the "share of political power allocated to a citizen with income $\tilde{y} = y(i, \omega)$ in state ω in a political system with wealth bias $\tilde{\alpha} = \alpha(\omega)$" (from here on, the notations $\tilde{y}$ and $\tilde{\alpha}$ are used to denote real values of the functions y and α, resp.). The interpretation is that λ represents the explicit features (e.g., constitutionally specified voting rules) of the political system. On the other hand α captures the nebulous features of a political system that are intrinsically hard to observe directly (e.g., effect of lobbying on a senator's vote or of campaign contributions on an election cycle).[12] Accordingly, the outside observer is assumed to know the function λ, but does not observe the bias function α. This isolates α as the object of interest, consistent

[11] We do require, and later verify, that the class of U satisfying (A1)-(A3) is nonempty.

[12] We point out, however, that our "Anything Goes" result (Theorem 1) does not depend on the observer having any knowledge of λ other than that it satisfies the axioms that follow.

with the main focus of the paper. The class of λ is defined by three axioms.

(B1) (Normalization) For each ω and $\tilde{\alpha}$, $\displaystyle\int_{y(0,\omega)}^{y(1,\omega)} \lambda(\tilde{y},\tilde{\alpha},\omega)\; dh(\tilde{y},\omega) \;=\; 1.$[13]

(B2) (Income Monotonicity). The function λ is assumed to be increasing in income level $\tilde{y}$ if $\tilde{\alpha} > 0$, decreasing in income if $\tilde{\alpha} < 0$; constant across income levels if $\tilde{\alpha} = 0$.

(B3) (Strict Single Crossing with Vanishing Tails) For each fixed ω, the function $\lambda(\tilde{y},\tilde{\alpha},\omega)$ satisfies strict single crossing in $(\tilde{\alpha};\tilde{y})$ with $\lim_{\tilde{\alpha}\to+\infty} \lambda(\tilde{y},\tilde{\alpha},\omega) = 0 \;\; \forall \tilde{y} < y(1,\omega)$, and $\lim_{\tilde{\alpha}\to-\infty} \lambda(\tilde{y},\tilde{\alpha},\omega) = 0 \;\; \forall \tilde{y} > y(0,\omega)$.

A simple example satisfying all the axioms is given by the vote share function,

$$\lambda(\tilde{y},\tilde{\alpha},\omega) \;=\; \frac{\tilde{y}^{\tilde{\alpha}}}{\displaystyle\int_{y(0,\omega)}^{y(1,\omega)} x^{\tilde{\alpha}} dh(x,\omega)} \tag{1}$$

In this case $\tilde{\alpha}$ exponentially weights wealth. One can then interpret $1 - \tilde{\alpha}$ as the weight attached to equal vote share or *equal representation* in voting.[14] Notice that the form of λ in (1) satisfies (B3) since the log of λ satisfies strict single crossing.

More generally, λ can be any voting share system consistent with the axioms. Axiom (B1) implies that the composite function $\lambda(y(i,\omega),\alpha(\omega),\omega)$ is a density in i. Axiom (B2) asserts that the political power of a citizen is a function of his income $\tilde{y}$, and the direction taken by λ depends on the sign of $\tilde{\alpha}$. Political power is increasing in income if $\tilde{\alpha} > 0$, decreasing if $\tilde{\alpha} < 0$, and invariant to income if $\tilde{\alpha} = 0$. Hence, the value $\tilde{\alpha}$ can be thought of as a measure of the extent of wealth bias in state ω. When $\tilde{\alpha} = 0$, the polity may be said to be *unbiased* in the sense that each person's political weight in the distribution is invariant to income, hence all individuals are political equals. We will refer to $\tilde{\alpha} > 0$ as the case of an *elitist bias* since wealth is rewarded in the political system. The case of $\tilde{\alpha} < 0$ is referred to as a *populist bias* since political power is redistributed away from wealth. We allow that the function α can take values in the entire real line.

In the canonical example in (1), for instance, when $\tilde{\alpha} = 1$ then an individual who possesses twice as much income as another has twice as many votes, hence twice as much political power.

[13] Recall that $h(\tilde{y},\omega_t)$ is the distribution of types over incomes $\tilde{y}$ as implied by the income process $y(\cdot)$.

[14] To see this more transparently, write (1) as

$$\lambda(\tilde{y},\tilde{\alpha},\omega) \;=\; \frac{\tilde{y}^{\tilde{\alpha}} 1^{1-\tilde{\alpha}}}{\displaystyle\int_{y(0,\omega)}^{y(1,\omega)} x^{\tilde{\alpha}} 1^{1-\tilde{\alpha}} dh(x,\omega)}.$$

The cases where $|\tilde{\alpha}| > 1$ are particularly stark in this example since this indicates a distribution of power that disproportionately rewards the fringes of the distribution. Extreme inequality occurs in the limit as $|\tilde{\alpha}| \to \infty$.

To understand the role of Axiom (B3), we use the normalization in (B1) to define the distribution function

$$L^p(j; \alpha, \omega) = \int_0^j \lambda(y(i,\omega), \alpha(\omega), \omega)\, di. \tag{2}$$

We refer to the distribution L^P as a *The Political Lorenz curve* since it gives a simple measure of political inequality. It describes the proportion of *political power* held by the poorest $j\%$ of types in state ω. Political inequality, as measured by L^P, can then change over time for two reasons. First, it can change due to changes in the income distribution. Second, it can change due to "structural" changes as captured by changes in $\alpha(\omega)$. Axiom (B3) is the key assumption in guaranteeing monotonicity in these structural changes as shown in the Lemma:

Lemma 1 *For every $j \in (0,1)$ and each ω,*

$$L^P(j, \alpha_2, \omega) < L^P(j, \alpha_1, \omega) \quad \forall\ \alpha_1(\omega) < \alpha_2(\omega).$$

The proof is in the Appendix. Under the Lemma, the absolute value $|\alpha(\omega)|$ can be used to measure the intensity of the bias. Larger positive values correspond to greater elitism in the bias - greater political inequality with weight accorded to wealth. A more negative $\alpha(\omega)$ corresponds to greater populism — again greater political inequality but in reverse. The extra asymptotic conditions in (B3) guarantee that political inequality hits the extremes (power allocated entirely to the richest or the poorest) in the limit as $\alpha(\omega) \to \infty$ or $\to -\infty$.

Using (1) as the canonical example, the Political Lorenz curve can be compared to the standard, income Lorenz curve given by

$$L(j, \omega) = \frac{\int_0^j y(i,\omega)di}{\int_0^1 y(i,\omega)di}, \tag{3}$$

As is standard, L describes the proportion of income held by the lowest j citizen-types in state ω. Figure 1a displays the two Lorenz curves in the case where the Political Lorenz curve exhibits a "dampened" elitist bias. Specifically, $0 < \alpha(\omega) < 1$, meaning that wealthier individuals have greater political weight than do poorer individuals, however, their increased weight is smaller than their weight in the income distribution. Political inequality therefore lies somewhere between income inequality and full equality. Figure 1b displays the Political Lorenz curve when $\alpha(\omega) > 1$. In that case the elitist bias is more pronounced, with political inequality that exceeds income inequality in the degree that the wealthy are accorded power. Note that the two curves coincide in the case where $\alpha(\omega) = 1$. Figure 2 illustrates the case of a populist bias, i.e., $\alpha(\omega) < 0$. Most theories we are aware of predict an elitist bias if any. Nevertheless, it does not seem sensible to rule out the $\alpha(\omega) < 0$ case, *a priori*.

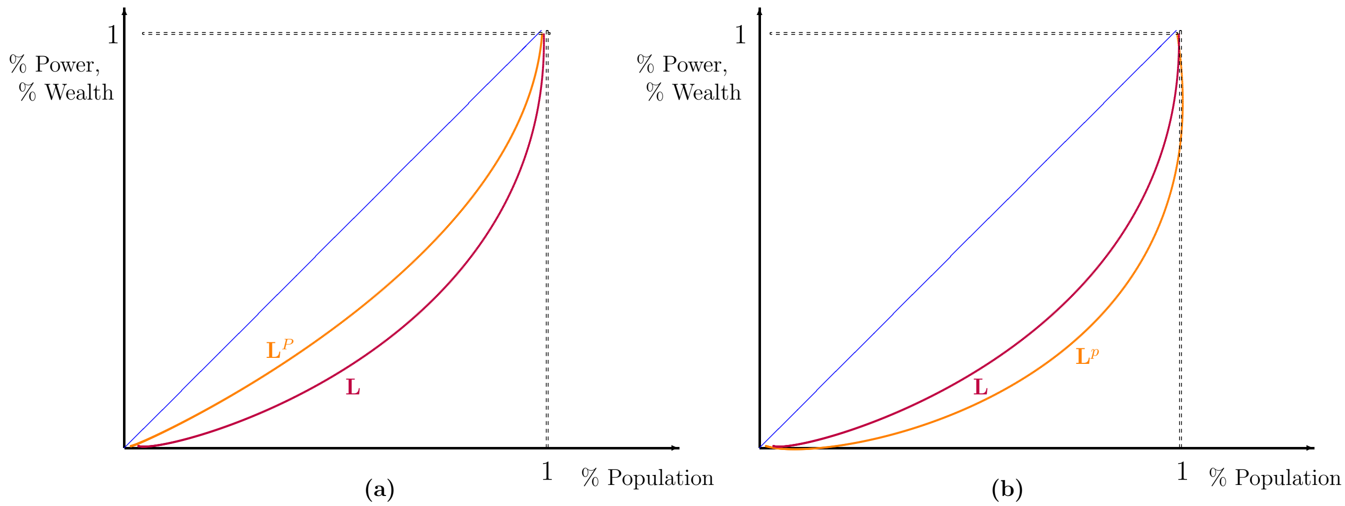

Figure 1: Political Lorenz Curves with Elitist Bias. (a) exhibits dampened bias. (b) exhibits pronounced bias.

3.2 Rationalizing Policy Data

Political Lorenz curves have a very simple interpretation. Suppose that policies are determined by some unspecified pairwise voting process. Each time a vote is taken, $\lambda(y(i,\omega),\alpha(\omega),\omega)$ is i's endowment of *vote share* in state ω. Policies are then determined by weighted majority voting where each individual's vote is weighted by his vote share. In the unbiased case ($\alpha(\omega)=0$), policies are determined by a simple majority vote.

Definition 1 Given a policy rule Ψ, a policy a is an α-*Weighted Majority Winner (WMW)* in state ω under admissible profile U if, for all policies $\hat{a}$,

$$\int_{i\in\{j:\,U(j,\omega,a;\Psi)\geq U(j,\omega,\hat{a};\Psi)\}} \lambda(y(i,\omega),\alpha(\omega),\omega)\;di \quad\geq\quad 1/2$$

In other words, an α-weighted majority winner, or α-WMW, is a policy that survives against all others in a majority vote when each type i is allocated $\lambda(y(i,\omega),\alpha(\omega),\omega)$ votes and the preference profile is given by U.

The unknown object of concern to the observer is the bias function α. If the preference profile U were known precisely to the outside observer, then α could be inferred precisely from observed policies that are generated from α (via the weighting function λ). But because U is not known, it is natural to ask whether observed policies might be "rationalized" by a weighting function α under some admissible preference profile U. Formally,

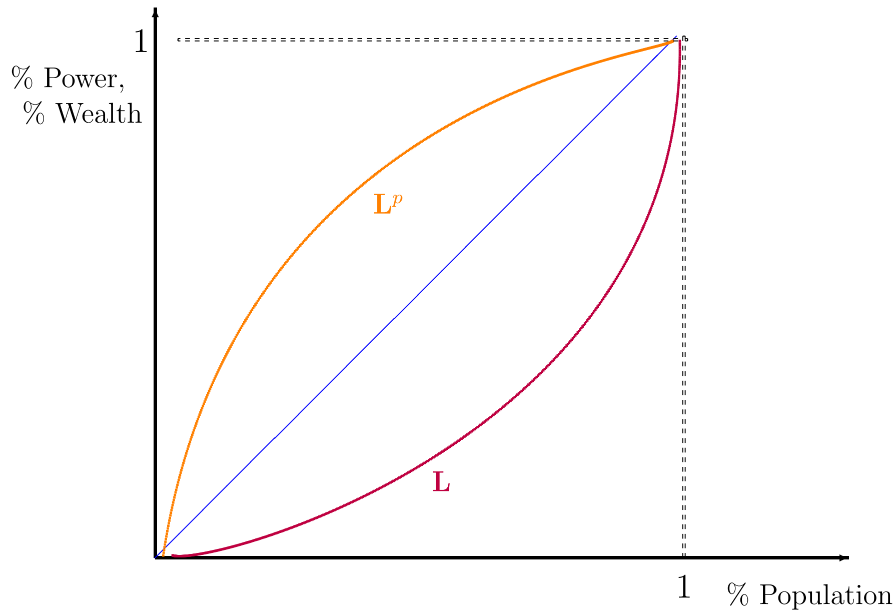

Figure 2: Political Lorenz Curve with Populist Bias

Definition 2 A weighting function α *rationalizes the observed policy data* $\{a_t\}_{t=1}^T$ if there exists an admissible profile U and a policy rule Ψ consistent with the data such that for all ω, $\Psi(\omega)$ is an α-weighted majority winner under U.

In words, α rationalizes $\{a_t\}_{t=1}^T$ if there is some policy rule Ψ consistent with data that can be produced by a political system with weighting function α. To make the connection with a given policy rule and a given profile explicit: we will sometimes refer to α as having *rationalized the data with rule* Ψ *under profile* U.[15]

The problem of figuring out which, if any, α rationalizes the data is made easier by applying a modified, dynamic version of the Median Voter Theorem. In this case, the "pivotal voter" is the weighted median type $i = \mu(\omega, \alpha)$ in the income distribution that implicitly solves

$$L^P(\mu(\omega, \alpha), \alpha, \omega) \; = \; \frac{1}{2} \tag{4}$$

The determination of $\mu(\omega, \alpha)$ is shown in Figure 3 for a particular $\alpha(\omega) > 0$. The following Lemma shows that the citizen-type $i = \mu(\omega, \alpha)$ is indeed pivotal.

Lemma 2 *Let* U *be an admissible profile. Then* α *rationalizes policy data with rule* Ψ *under*

[15]Alternatively, one could refer to a policy data as being "rationalized by the triple (α, Ψ, U)". This turns out to be notationally cumbersome and detracts from the main focus on bias weight α.

profile U if and only if

$$\Psi(\omega) \in \arg\max_{a \in A} U(\mu(\omega,\alpha),\omega,a;\Psi), \; \forall \; \omega.$$

The Lemma follows from the single crossing property (A2) on preferences, however, for completeness the proof is in the Appendix. Clearly, since α is latent, the pivotal type is as well. However, it will sometimes prove more convenient to consider inference over $\mu(\omega,\alpha)$ rather than over α directly.

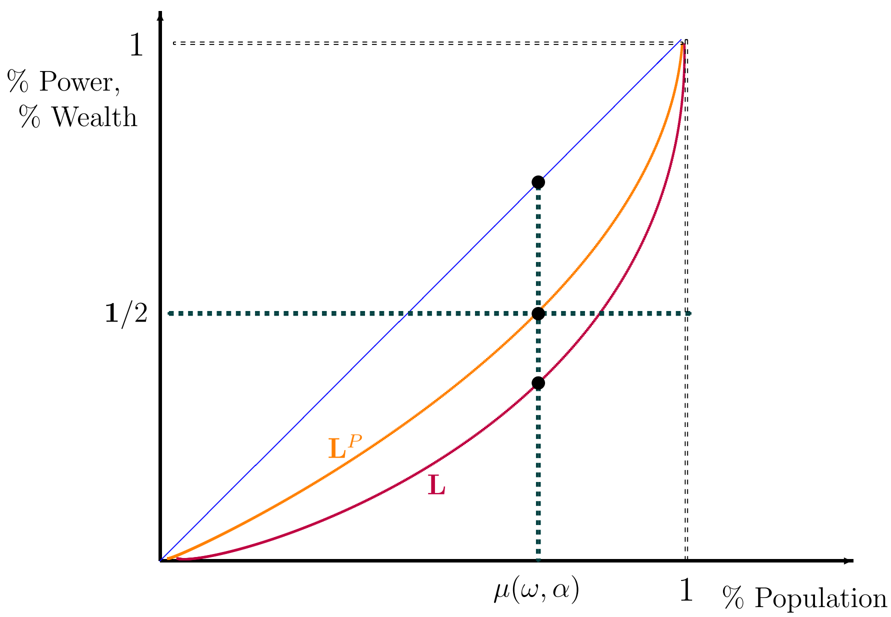

Figure 3: Identifying a Pivotal Voter under an Elitist Bias

It can also be shown from Lemma 1 that $\mu(\omega,\alpha_1) > \mu(\omega,\alpha_2)$ whenever $\alpha_1(\omega) > \alpha_2(\omega)$. To see this in a simple parametric example, let income process be given by,

$$y\,(i,\omega) = \exp(i\omega). \tag{5}$$

Here, one can interpret i as the production efficiency of type i and ω as the public capital stock. It is easy to see that $y\,(i,\omega)$ is increasing in (i,ω). In addition, income inequality increases in ω. Combining (5) with the vote share function in (1) yields

$$\lambda\,(y(i,\omega),\alpha(\omega),\omega) \;=\; \frac{y\,(i,\omega)^{\alpha(\omega)}}{\int_j y\,(j,\omega)^{\alpha(\omega)}\,dj} \;=\; \frac{\exp(\alpha(\omega)\omega i)}{\int_j \exp(\alpha(\omega)\omega j)dj}.$$

This implies a Political Lorenz curve given by

$$L^p\left(j, \alpha, \omega\right) \;=\; \int_0^j \lambda\left(y(i, \omega), \alpha(\omega), \omega\right) di \;=\; \frac{exp(\alpha(\omega)\omega j) - 1}{exp(\alpha(\omega)\omega) - 1}.$$

and, consequently, a pivotal voter function $\mu(\omega, \alpha)$ given by

$$\mu(\omega, \alpha) \;=\; \frac{log[\frac{1}{2}(exp(\alpha(\omega)\omega) + 1)]}{\alpha(\omega)\omega}$$

4 Results

Starting with Markov consistent observations, one could ask first whether the observations could have been produced by any weighting function, and second whether the observations could have been produced by a particular type of weighting function. The first of our results answers both questions at once.

4.1 An "Anything Goes" Theorem

Theorem 1 *Let* $\{a_t\}_{t=1}^T$ *be any policy data and let* α *be any weighting function. Then* α *rationalizes* $\{a_t\}_{t=1}^T$.

According to Theorem 1, without specific information about preference orderings, the policy data does not tell us anything about political inequality, whether it exists or whether its magnitude is large. Since, among all other α, the unbiased polity $\alpha(\omega) = 0\ \forall\omega$ can also rationalize policy data, we cannot rule it out.

Proof. Let Ψ be a policy rule consistent with the policy data. Consider any long run payoff U of the form

$$U\left(i, \omega, a; \Psi\right) = -\frac{1}{2}\left(a^2 - \Psi^2\left(\omega\right)\right) + \widetilde{\Psi}\left(i, \omega\right)\left(a - \Psi\left(\omega\right)\right), \tag{6}$$

where $\widetilde{\Psi}\left(i, \omega\right)$ is continuous and weakly increasing in i for every $\omega \in \Omega$, and satisfies $\widetilde{\Psi}(\mu(\omega, \alpha), \omega) = \Psi(\omega)$. Notice that Equation (6) defines a non-empty set of U for any given $\alpha(\omega)$ and any $\Psi(\omega)$, a concrete example being $\widetilde{\Psi}(i, \omega) = i - \mu(\omega, \alpha) + \Psi(\omega)$. We are going to show that for any $\Psi(\omega)$ consistent with the data, any α rationalizes the data with $\Psi(\omega)$ under any U defined in (6).

We first verify that U is admissible. For (A1), observe that U is continuous in i and strictly concave in a (hence single peaked). From the weak increasing property of $\widetilde{\Psi}$ in i, U is single crossing in $(a;i)$, as required in (A2). Given the monotonicity of $y(i,\omega)$ in i, there is an inverse function $h(\omega,y)=i$ such that h is increasing in y. To verify (A3), we now find the flow payoff u as the difference:

$$\begin{aligned} u(\omega,y,a) &= U(h(\omega,y),\omega,a) \; - \; \delta U(h(\omega,y),Q(\omega,a),\Psi(Q(\omega,a)) \;) \\ &= \; - \; \tfrac{1}{2}[a^2-(\Psi(\omega))^2] \; + \; \tilde{\Psi}(h(\omega,y),\omega)[a-\Psi(\omega)]. \end{aligned} \tag{7}$$

Therefore, U as defined in (6) is admissible.

From the first order conditions, $\widetilde{\Psi}(i,\omega)$ is the preferred policy choice for type i under ω. Applying Lemma 2, α rationalizes the data under U in (6) if and only if there exists a $\widetilde{\Psi}$ and Ψ such that $\widetilde{\Psi}(\mu(\omega,\alpha),\omega)=\Psi(\omega)$. But this is true by construction. This finishes the proof. ∎

At this stage there are two possible ways one could rule out certain bias weights. First, one could add direct information about specific binary rankings. Such information could come, for instance, from polls. We consider this option in Section 5. Before proceeding with that option, however, we explore a second option: that the class of admissible profiles considered so far is "too large" and must be pared down.

4.2 A Monotone Comparative Statics Restriction

Is there a sensible subset of admissible profiles such that a given α *cannot* rationalize policy data in this class of preferences? One possible, though by no means only, requirement is that an individual's bliss rule is monotone, increasing in the value of the state. Formally, consider

(A4) U satisfies single crossing in the pair $(a\;;\omega)$ for each i.

Assumption (A4) implies that every type's most preferred policy is (weakly) monotone in the state. This monotonicity restriction is fairly common when the policy is a complementary input in the production process. It is straightforward to show that if U is an admissible profile that also satisfies (A4) then the policy $\tilde{\Psi}(i,\omega_t)$ that maximizes $U(i,\omega_t,a_t;\Psi)$ is weakly increasing in both i and ω_t.

A weighting function α will be said to *sc-rationalize* the policy data $\{a_t\}$ if it rationalizes this data under an admissible profile U that satisfies (A4). The following result shows that sc-rationalization imposes meaningful restrictions on α.

Theorem 2 *Let $\{a_t\}_{t=1}^{T}$ be any policy data. Then:*

1. *There exists an α that sc-rationalizes the data.*

2. *Any given α sc-rationalizes policy data $\{a_t\}$ if and only if for each pair of observed states ω_t, ω_τ with $\omega_t > \omega_\tau$,*

$$a_t < a_\tau \implies \mu(\omega_t, \alpha) < \mu(\omega_\tau, \alpha) \tag{8}$$

The necessary condition in Part 2 is straightforward. Suppose α sc-rationalizes $\{a_t\}$ with Ψ under some admissible U satisfying (A4) as hypothesized. Let $\tilde{\Psi}(i, \omega)$ maximize $U(i, \omega, a;\ \Psi)$. Then by (A4), $\tilde{\Psi}(i, \omega)$ is weakly increasing in ω and in i. Since $\Psi(\omega) = \tilde{\Psi}(\mu(\omega, \alpha), \omega)$ for all ω, (8) must follow. A consequence of (8) is that any policy data increasing in the observed state can be rationalized by any α. However, if the data is ever decreasing in the state, then certain α may not rationalize the data. For instance:

Corollary *Let $\{a_t\}$ be any policy data such that the observed policies decrease whenever the state increases. Then the unbiased weighting function, $\alpha(\omega) = 0$ for all ω, does not sc-rationalize $\{a_t\}$.*

The sufficiency proof of Theorem 2 requires a constructive argument just as in Theorem 1. As in that result, we construct a payoff function of the form in (6). The function $\tilde{\Psi}(i, \omega)$ must be constructed to satisfy the preference axioms while, at the same time, match the policy data on the observed path whenever type i is the pivotal voter, $\mu(\omega, \alpha)$. Unlike the Theorem 1 proof, however, the function $\tilde{\Psi}(i, \omega)$ must also be weakly increasing in both ω and i in order to satisfy the additional axiom, (A4). Though this sounds simple, the construction is complicated by the fact that while $\tilde{\Psi}$ must be increasing in the state, the actual policy data may not be. To overcome this, we specify a recursive algorithm that exploits the natural monotonicity of the data in i_t — as required by the hypothesis in (8). The formal argument is left to the Appendix.

5 Revealed Political Power and the Power of Polls

External information about policy preferences often exists in the form of polls. This section examines how simple aggregate data from polls might reveal information about political bias.

Consider the following scenario. Each period t, a poll is taken in which citizens are asked to compare the actual policy choice $a_t = \Psi(\omega_t)$ to some small collection of fixed alternatives in the feasible policy set A. Typically, these alternatives are some much discussed policy alternatives, always on the table but not necessarily adopted.

5.1 Rationalizing Policy and Polling Data

To begin, we examine the case of two anonymous binary polls that ask individuals to rank a_t against each of two alternatives, $\bar{a}$ and $\underline{a}$ such that $\underline{a} < a_t < \bar{a}$. Policy alternative $\underline{a}$ can be thought of as the "left wing" alternative to the chosen policy, $\bar{a}$ the "right-wing" alternative. The analysis is extended later on to allow for any number of polls. For tractability, these polls are assumed to be accurate in the sense that the sampling error is ignored.[16] The poll data is summarized by a path $\{p_t, q_t\}_{t=1}^T$ such that at each date $t = 1, \ldots, T$, p_t and q_t represent the fractions of the population that weakly prefer the weighted-majority winner a_t to the alternatives $\bar{a}$ and $\underline{a}$, respectively, in state ω_t.

Since underlying profile U that generates the poll data is itself unobservable, the poll data must be consistent with both U and the observable policy data.

Definition 3 A weighting function α *rationalizes both policy data* $\{a_t\}_{t=1}^T$ *and poll data* $\{p_t, q_t\}_{t=1}^T$ if there exists an admissible U and a policy rule Ψ consistent with the data such that

(i) for each $\omega \in \Omega$, the policy rule $\Psi(\omega)$ is an α-weighted majority winner under payoff function U, and

(ii) for each $t = 1, \ldots T$, U satisfies

$$\begin{aligned} p_t &= |\{i : \ U(i, \omega_t, a_t; \Psi) \geq U(i, \omega_t, \bar{a}; \Psi)\}|, \quad \textit{and} \\ q_t &= |\{i : \ U(i, \omega_t, a_t; \Psi) \geq U(i, \omega_t, \underline{a}; \Psi)\}| \end{aligned} \tag{9}$$

Part (i) is the policy-consistency requirement as before. Part (ii) is a poll-consistency requirement. It requires that the admissible profile U must reflect preferences that can generate the observed poll data consistent with Ψ.

What type of weighting function α is consistent with the policy and poll data? To derive necessary conditions, we apply properties of the preference class. From the single crossing property (A2), it follows that at each date t the poorest fraction p_t weakly prefer the observed policy $a_t = \Psi(\omega_t)$ to alternative $\bar{a}$, and the richest fraction q_t weakly prefer a_t to alternative $\underline{a}$.

$$\begin{aligned} \{i : \ U(i, \omega_t, a_t; \Psi) \geq U(i, \omega_t, \bar{a}; \Psi)\} &= [0, p_t], \quad \textit{and} \\ \{i : \ U(i, \omega_t, a_t; \Psi) \geq U(i, \omega_t, \underline{a}; \Psi)\} &= [1 - q_t, 1]. \end{aligned}$$

In fact, the assumption that a_t must be a weighted-majority winner implies certain restrictions on p_t and q_t.

[16]Though obviously unrealistic, this is in keeping with the benchmark case of all data having been produced by a deterministic politico-economic system.

First, observe that

$$1 - q_t < p_t \tag{10}$$

If this were not the case then a nonempty interval $[p_t, 1 - q_t]$ exists, consisting of types that weakly prefer both the smallest policy, $\underline{a}$, and the largest, $\bar{a}$, to a_t. This, in turn, is a contradiction of the single peakedness assumption (A1) on U.

Second, notice that for any $0 < p_t < 1$, the pivotal voter $\mu(\omega_t, \alpha)$ must be smaller than p_t. If, for instance, $\mu(\omega_t, \alpha) > p_t$, then Lemma 2 is violated. In that case the fraction $(p_t, 1]$ who prefer the right-wing alternative $\bar{a}$ would exceed half the weighted vote share. This would place the supporters of $\bar{a}$ in a position to have vetoed a_t, in which case a_t could not have been a WMW. If $\mu(\omega_t, \alpha) = p_t$, then by continuity of U in i (Assumption (A1)), this pivotal voter would be indifferent between a_t and $\bar{a}$, thus violating single peakedness. Consequently, we must have $\mu(\omega_t, \alpha) < p_t$. Finally, if $p_t = 1$, then $\mu(\omega_t, \alpha) < p_t$ also holds by construction.

Similar arguments establish that $1 - q_t < \mu(\omega_t, \alpha)$ for all $0 < q_t \leq 1$. Together, these facts establish the necessary conditions in the following result.

Theorem 3 *Let $\{a_t\}_{t=1}^T$ be any policy data and $\{p_t, q_t\}_{t=1}^T$ be any polling data. Then:*

1. *There exists an α that rationalizes the data if and only if $p_t > 1 - q_t$ for all t.*
2. *Any given α rationalizes $\{a_t\}_{t=1}^T$ and $\{p_t, q_t\}_{t=1}^T$ if and only if*

$$1 - q_t \; < \; \mu(\omega_t, \alpha) \; < \; p_t, \quad \forall \; t = 1, \ldots, T \tag{11}$$

The "sufficiency" argument in part 2, i.e., the construction of a U under which Ψ and $\{p_t, q_t\}$ are rationalized, appears in the Appendix.

Equation (11) defines implicit upper and lower bounds on the set of rationalizing weight $\alpha(\omega_t)$. These bounds can be made explicit as follows. Notice first that the pivotal voter function μ is itself defined implicitly by (4). Since L^P is decreasing in the weight $\alpha(\omega)$ (holding ω fixed), the pivotal function μ is invertible in the value $\alpha(\omega)$. Let $M(j, \omega) = \tilde{\alpha}$ denote this *inverse pivotal function.* It is readily verified that M is increasing in j, and by definition, $\alpha(\omega) = M(\mu(\omega, \alpha), \omega)$ must hold. Equation (11) then implies

$$M(1 - q_t, \omega_t) \; < \; \alpha(\omega_t) \; < \; M(p_t, \omega_t), \quad \forall \; t = 1, \ldots, T \tag{12}$$

In other words, the inverse pivotal function is used to define a band of rationalizing weights α. We refer to this interval $(M(1 - q_t, \omega_t), M(p_t, \omega_t))$ as the *bias band.* Figure 4 expresses a graph of M and the bias band when λ has the canonical form given by Equation (1). The bounds of the band are displayed on the vertical axis. In the graph, the range of bias band includes 0, the unbiased weight. It also includes a subinterval of elitist biases, as well as a subinterval of populist ones. Using the inverse pivotal function M, elitist or populist bias can often be identified.

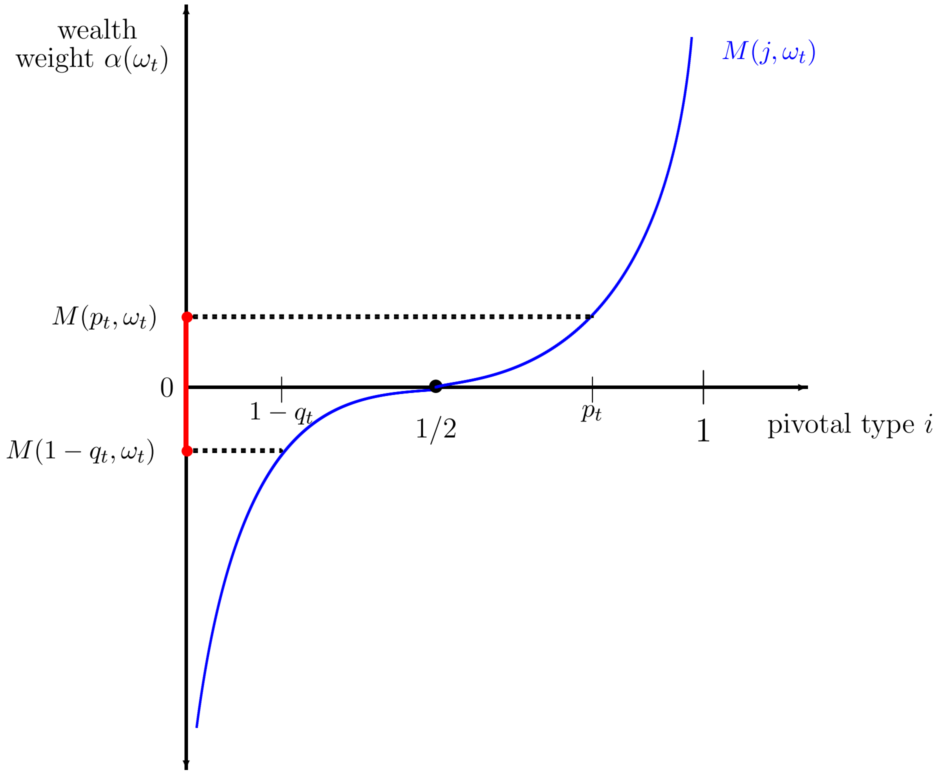

Figure 4: Bias Band and Bounding Function

Theorem 4 *Suppose that α rationalizes the policy data $\{a_t\}$ and poll data $\{p_t, q_t\}$. Then for each $t = 1, \ldots, T$,*

(i) $M(1-q_t, \omega_t) \geq 0$ *whenever* $q_t < 1/2$.

(ii) $M(p_t, \omega_t) \leq 0$ *whenever* $p_t \leq 1/2$.

Part (i) identifies an elitist bias. Part (ii) then identifies a populist bias.

In addition to state-by-state bounds implied by polling data, the polls may also impose dynamic restrictions. Observe that the Political Lorenz curve can change over time for two reasons. First, changes in the state directly affect income inequality, and power is wealth-weighted. Hence, political inequality changes with changes in the income distribution. Second, Political Lorenz curve changes as the bias weight $\alpha(\omega_t)$ changes, holding fixed the income process. Generally, the two causal effects are hard to decouple. Sometimes, spatial information can be used to sort them out. One simple extreme case occurs when $1 - q_{t+1} > p_t$. That is, the individuals who prefer $\underline{a}$ to a_{t+1} in state ω_{t+1} is more numerous than those who prefer a_t to $\bar{a}$ in state ω_t. This condition implies that the distribution of ideal points has, roughly speaking, shifted to the left. It is straightforward to show

Theorem 5 *Suppose α rationalizes $\{a_t\}$ and $\{p_t, q_t\}$ and income inequality remains stable. Then:*

(i) $\alpha(\omega_{t+1}) > \alpha(\omega_t)$, i.e., the bias in $t+1$ is unambiguously more elitist than in t if $p_t < 1 - q_{t+1}$.

(ii) $\alpha(\omega_{t+1}) < \alpha(\omega_t)$, i.e., the bias in $t+1$ is unambiguously more populist than in t if $p_{t+1} < 1 - q_t$.

In words, the bias becomes more elitist over time if there are individuals who prefer both $\bar{a}$ in date t and $\underline{a}$ in $t+1$. This means that there are right-wing dissidents who oppose the chosen policy in t who later become left-wing dissidents in $t+1$. Similarly, the bias becomes more populist if there are individuals who prefer both $\underline{a}$ in date t and $\bar{a}$ in $t+1$. Hence, left-wing dissidents in t become right-wing dissidents in $t+1$.

Proof. Consider (i). Suppose $p_t < 1 - q_{t+1}$. Then the date t bias band, given by $(M(1-q_t, \omega_t), M(p_t, \omega_t))$ lies entirely below the date $t+1$ bias band $(M(1-q_{t+1}, \omega_{t+1}), M(p_{t+1}, \omega_{t+1}))$. By Theorem 3, $\alpha(\omega_{t+1}) > \alpha(\omega_t)$. In other words, the bias becomes more elitist in $t+1$.

Now consider (ii). Suppose $p_{t+1} < 1 - q_t$. Using the same reasoning as in (i), this implies that the date t bias band lies entirely *above* the date $t+1$ bias band. Thus by Theorem 3, $\alpha(\omega_{t+1}) < \alpha(\omega_t)$ so that the bias becomes more populist in $t+1$. ∎

5.2 Income Inequality and Political Inequality

In this section, we restrict the vote share function λ to the canonical form in (1), in which case $\alpha(\omega)$ exponentially weights income. Next, consider an income process y that has monotone log differences in the pair (i, ω). That is, for any pair of states, ω and $\hat{\omega}$ the difference $\log y(i, \omega) - \log y(i, \hat{\omega})$ is either increasing in i, or decreasing in i. Many common income processes, including the earlier example in (5) satisfy this condition.

These restrictions allow us to make a direct comparison between income and political inequality.

Lemma 3 *Suppose y has monotone log differences in the pair (i, ω). Then for any pair of states ω_1 and ω_2 such that $\alpha(\omega_1) = \alpha(\omega_2) = \tilde{\alpha}$, $L(j, \omega_1) > L(j, \omega_2)$ for all $j \in (0, 1)$, iff $L^P(j, \alpha, \omega_1) > (<)\ L^P(j, \alpha, \omega_2)$ for all $j \in (0, 1)$ and $\tilde{\alpha} > (<) 0$.*

In words, when under monotone log differences, political inequality increases whenever income inequality does so as well. Note that when the bias is populist, increasing political inequality rewards poorer types.

The reasoning is straightforward. Both increased income and political inequality are statement about first-order stochastic orderings. In either case, a distribution under, say, ω_2 first-order dominates a distribution under ω_1 if the likelihood ratio is increasing. The log-likelihood ratio of L is

$$\log y(i, \omega_2) - \log y(i, \omega_1) - \log \left(\frac{\int_0^1 y(s, \omega_2)\, ds}{\int_0^1 y(s, \omega_1)\, ds} \right),$$

whereas the log-likelihood ratio of L^P is

$$\widetilde{\alpha} \left[\log y(i, \omega_2) - \log y(i, \omega_1) \right] - \log \left(\frac{\int_0^1 y(s, \omega_2)^{\widetilde{\alpha}}\, ds}{\int_0^1 y(s, \omega_1)^{\widetilde{\alpha}}\, ds} \right).$$

Now suppose that the log difference of y is increasing in i. Then both likelihood ratios are increasing if $\widetilde{\alpha} > 0$. Similarly, the likelihood ratio for L is increasing, and that for L^P is decreasing if $\widetilde{\alpha} < 0$. In turn, the following result is immediate from Lemma 1 and Lemma 3.

Theorem 6 *Let ω_1 and ω_2 be any two states that satisfy $L(j, \omega_2) < L(j, \omega_1)$ for all j, i.e., income inequality is larger in state ω_2. Then $|M(j, \omega_2)| < |M(j, \omega_1)|$ for all $j \neq 1/2$. In particular, if* 0 *(the unbiased weight) belongs to the band, then larger income inequality reduces the size of the band around 0.*

According to the result, increases in income inequality have an equalizing effect politically, other things equal. When 0 is an admissible weight, then the band shrinks around it. If the band is entirely above 0 (elitism) then it moves closer to 0. Intuitively, this is not surprising since in that case, the pro-wealth bias must be lower to have off-set the greater income inequality. Somewhat more surprising is the fact that when the band is entirely below 0 (populism), greater inequality moves the band closer to 0 as well. In other words the band becomes less populist implying that wealthier individuals receive increased political weight from the bias in addition to increased weight from income alone. Why? Because with a populist system, political inequality is a negatively related to income inequality. Hence, holding the bias weight constant, an increase in relative income of the top 10% translates into an weighted decrease in this group's political power. The bias weight must therefore increase to offset this fall in political power due to income change. This dual effect of greater income inequality is displayed in Figure 5.

5.3 Many Polls

This subsection extends the analysis to an arbitrary number of polls taken each period. It's worth noting, first, that an added poll that compares the left wing alternative $\underline{a}$ to right wing one $\bar{a}$ adds no restrictions directly on α since it doesn't involve the political system. The poll does, however, place additional restrictions on the admissible U. To see this, let r_t be the

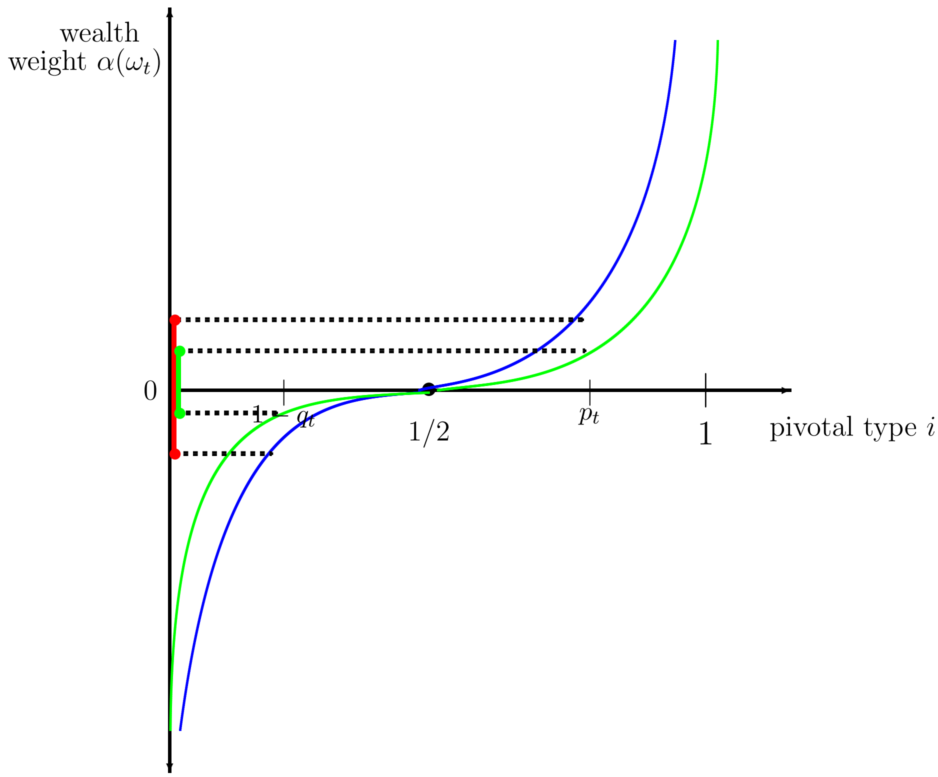

Figure 5: Shrinking Bias Band with Increased Inequality

support rate for $\underline{a}$ against $\bar{a}$. Then, following the previous discussion, the policy and poll data can be rationalized only if $1 - q_t < r_t < p_t$. If this were not the case, then no single peaked preference could have generated the data.

Consider, next, an arbitrary number N of polls taken each period, each pitting the observed a_t against an alternative policy. Let $a^1 < a^2 < \ldots < a^N$ denote alternatives specified in the poll data. Assume $0 < p_t^n < 1$ is the support rate that weakly favors the observed policy a_t against policy a^n. Using notational convention $a^{N+1} = \max a$, let $p^{N+1} = 1$. Then let $n_t^* = \min\{n : \ a^n \geq a_t, \ 1 \leq n \leq N+1\}$.

Theorem 7 *Let $\{a_t\}$ be any policy data and $\{p_t^n\}_{t=1,n=1}^{T,N}$ any arbitrary polling data. Then:*

1. *There exists an α that rationalizes the policy and poll data if and only if for each t,*

$$1 - p_t^1 < \ldots < \ 1 - p_t^{n_t^*-2} \ < \ 1 - p_t^{n_t^*-1} \ < p_t^{n_t^*} \ < \ p_t^{n_t^*+1} \ < \ldots < p_t^N \tag{13}$$

2. *Any given α rationalizes the policy and poll data if and only if*

$$1 - p_t^1 < \ldots < 1 - p_t^{n_t^*-2} \ < \ 1 - p_t^{n_t^*-1} \ < \ \mu(\omega_t, \alpha) \ < \ p_t^{n_t^*} \ < \ p_t^{n_t^*+1} \ < \ldots < p_t^N \tag{14}$$

The Theorem is a straightforward extension of Theorem 3. Using the inverse pivotal function M, the bounds on α are given by

$$M(1-p_t^{n_t^*-1},\omega_t) \ < \ \alpha(\omega_t) \ < \ M(p_t^{n_t^*},\omega_t)$$

Hence, any direct inference of α depends only on the comparisons between the observed policy and those closest to it among the alternatives. The remaining comparisons are, in a sense, redundant (though they do further restrict the U).

6 Summary and Extensions

This paper adapts ideas from revealed preference theory to understand political bias. To assess the bias, we formulate a theory of inference based on an outside observer's direct view of policy rather than on indirect measures such as political participation. The theory associates political bias with the weights on a system of wealth-weighted majority voting.

Given fairly standard assumptions ensuring each admissible preference profile admits a weighted majority winner, every weighted system is shown to rationalize every possible policy path. Further restrictions on preferences can rule out certain weighted systems. The introduction of polling data rules out "extreme" weighting systems, by imposing upper and lower bounds on the magnitude of the bias.

It turns out that some, though not all the rationalizing characterizations hold in an even stronger sense. Say that a weighting function α *strongly rationalizes the policy data* if for *all* policy rules Ψ consistent with the data, there exists an admissible profile U such that for all ω, $\Psi(\omega)$ is an α-weighted majority winner under U. In other words, it is not enough that one can find some policy rule consistent with data, the rationalizing weights have to be able to reproduce all such policy rules.

A simple inspection of the proofs reveal that the conclusions of all the theorems except one hold when "strongly rationalize" replaces "rationalize" in the results. The lone exception is Theorem 2. Its proof makes use of an algorithm (see Appendix) which restricts off-path behavior in a particular way.

As for limitations, the main hurdle in our view is the restriction to one dimensional policies and states. The dimensionality restriction, together with single crossing ensure existence of a majority winner. If policies and states are multi-dimensional, then the single crossing condition on the natural (Euclidian) order is no longer sufficient to ensure majority voting outcomes. At this point one's options are limited. One option is to use a common generalization of (A2), known as "order restrictedness", due to Rothstein (1990). Order restricted preferences are those for which there exists *some* order on the policy space A (other than, presumably, the Euclidian order) under which preferences are single crossing. Under order restricted preferences, wealth-weighted majority winner always exist. Because this is a fairly direct extension, we omit details.

A second and more challenging option is to drop all assumptions that guarantee existence of weighted majority voting. Even in this case, it is possible to articulate a well defined theory, albeit one with few known implications. A weaker equilibrium notion, for instance, the *weighted minmax majority winner (WMMW)* can always be shown to exist. Roughly, WMMW's are policies that garner more support than any other when policies are pitted against their most popular alternatives.[17] It's straightforward to show that the set of WMMW is always nonempty, and coincides with the set of WMW whenever the latter is nonempty. A drawback of this generalization is that since policies are not necessarily well ordered, it is not clear how changes in observed policy map into wealth distribution. Consequently, it is hard to see how meaningful inference is possible at this level of generality. We have little to say about it at this point, and so we leave it for future consideration.

7 Appendix

Proof of Lemma 1 Let $f(i,\alpha,\omega) = \lambda(y(i,\omega),\alpha,\omega)$ and define

$$D(x) = L(x,\alpha_2,\omega) - L(x,\alpha_1,\omega) = \int_0^x (f(i,\alpha_2,\omega) - f(i,\alpha_1,\omega))\,di.$$

From strict single crossing property, $f(i,\alpha_2,\omega) - f(i,\alpha_1,\omega) \geq 0$ implies $f(j,\alpha_2,\omega) - f(j,\alpha_1,\omega) > 0$ for every $j > i$. By definition, $D(0) = 0$ and $D(1) = 0$. As a result, it cannot be the case that $f(i,\alpha_2,\omega) - f(i,\alpha_1,\omega) > 0$ or $f(i,\alpha_2,\omega) - f(i,\alpha_1,\omega) < 0$ for almost all $i \in (0,1)$. Consequently, as a function of i, $f(i,\alpha_2,\omega) - f(i,\alpha_1,\omega)$ crosses zero exactly once and from below. This implies that $L(x,\alpha_2,\omega) < L(x,\alpha_1,\omega)$ for every $x \in (0,1)$.

Proof of Lemma 2 Fix an admissible profile U. The sufficiency ("if") part follows directly from the single crossing property (A2) (see Gans and Smart, 1996). For necessity ("only if"), suppose that α rationalizes the policy data with Ψ. Suppose further for at least one $\hat{\omega}$, $\Psi(\hat{\omega}) \notin \arg\max_{a \in A} U(\mu(\hat{\omega},\alpha),\hat{\omega},a;\Psi)$. Then there exists an $\hat{a}$ such that $U(\mu(\hat{\omega},\alpha),\hat{\omega},\hat{a};\Psi) > U(\mu(\hat{\omega},\alpha),\hat{\omega},\Psi(\hat{\omega});\Psi)$. From the single crossing property (A2) and from continuity of U in i (A1), more than half of voters strictly prefer $\hat{a}$ to $\Psi(\hat{\omega})$, contradicting the fact that $\Psi(\hat{\omega})$ is an α-WMW in state $\hat{\omega}$.

[17] Formally, a policy a is a *Weighted-Minmax Majority Winner (WMMW)* in state ω if

$$\int_{\{i:\, U(i,\omega,a;\Psi) \geq U(i,\omega,\hat{a};\Psi)\}} \lambda(y(i,\omega),\alpha(\omega),\omega)\; di \quad \geq \int_{\{i:\, U(i,\omega,a';\Psi) \geq U(i,\omega,\hat{a}';\Psi)\}} \lambda(y(i,\omega),\alpha(\omega),\omega)\; di$$

for all $\hat{a}$, a' and $\hat{a}'$.

Proof of Theorem 2.

Part 1. To prove Part 1, existence of a rationalizing α, we first suppose that Part 2 holds. Let $\{i_t\}_{t=1}^T$ be a sequence of positive real numbers such that for every $\omega_t > \omega_\tau$, $a_t < a_\tau$ implies $i_t < i_\tau$. In this case it suffices to show that there exists α such that $\mu(\omega_t, \alpha) = i_t$ for every ω_t. Because $\mu(\omega, \alpha)$ is continuous in α, we proceed to show this via a full range condition. That is, $\mu(\omega, \alpha) \to 1$ if $\alpha \to +\infty$; $\mu(\omega, \alpha) \to 0$ if $\alpha \to -\infty$. The argument, using Axiom (B2), is as follows. By definition, $L^P(j, \tilde{\alpha}, \omega) = \int_0^j \lambda(y(i, \omega), \tilde{\alpha}, \omega)\, di$.
(a) If $\lim\limits_{\tilde{\alpha}\to+\infty} \lambda(\tilde{y}, \tilde{\alpha}, \omega) = 0$, $\forall \tilde{y} < y(1, \omega)$, then $\lim\limits_{\tilde{\alpha}\to+\infty} \lambda(y(i, \omega), \tilde{\alpha}, \omega) = 0$, $\forall i < 1$. Hence $\lambda(y(i, \omega), \tilde{\alpha}, \omega)$ is a uniformly bounded function for every fixed $i < 1$ when $\tilde{\alpha}$ is large enough. For every $j < 1$, $\int_0^j \lim\limits_{\tilde{\alpha}\to+\infty} \lambda(y(i, \omega), \tilde{\alpha}, \omega)\, di = 0$. From Bounded Convergence Theorem on $[0, j]$, one can exchange lim and integral to obtain $\lim\limits_{\tilde{\alpha}\to+\infty} \int_0^j \lambda(y(i, \omega), \tilde{\alpha}, \omega)\, di = 0$, or $\lim\limits_{\tilde{\alpha}\to+\infty} L(j, \tilde{\alpha}, \omega) = 0$.
(b) If $\lim\limits_{\tilde{\alpha}\to-\infty} \lambda(\tilde{y}, \tilde{\alpha}, \omega) = 0$, $\forall \tilde{y} > y(0, \omega)$, then $\lim\limits_{\tilde{\alpha}\to-\infty} \lambda(y(i, \omega), \tilde{\alpha}, \omega) = 0$, $\forall i > 0$. Hence $\lambda(y(i, \omega), \tilde{\alpha}, \omega)$ is a uniformly bounded function for every fixed $i > 0$ when $\tilde{\alpha}$ is small enough. For every $j < 1$, $\int_{1-j}^1 \lim\limits_{\tilde{\alpha}\to-\infty} \lambda(y(i, \omega), \tilde{\alpha}, \omega)\, di = 0$. Again using the Bounded Convergence Theorem on $[1-j, 1]$, the lim and integral are exchanged to obtain $\lim\limits_{\tilde{\alpha}\to-\infty} \int_{1-j}^1 \lambda(y(i, \omega), \tilde{\alpha}, \omega)\, di = 0$, or $\lim\limits_{\tilde{\alpha}\to-\infty} (1 - L(j, \alpha, \omega)) = 0$ and hence $\lim\limits_{\tilde{\alpha}\to-\infty} L(j, \tilde{\alpha}, \omega) = 1$.

Part 2. We now establish Part 2. Since the necessary condition was proven in the paper, it remains "only" to show the sufficiency argument. Hence, fix any policy data and weighting function α that satisfy the implication in (8).

Now consider a policy rule $\Psi(\omega)$ consistent with the data and a payoff function U of the form in (6) given in the Proof of Theorem 1. In that construction, $\widetilde{\Psi}(i, \omega)$ is continuous and weakly increasing in i. The Theorem 1 proof goes on to show that U is admissible (satisfies (A1)-(A3)), and α rationalizes data using the policy function Ψ under payoff U. Notice that if $\widetilde{\Psi}(i, \omega)$ also happens to be weakly increasing in ω, then U would be admissible *and* satisfy (A4) as required for the present result.

To complete the proof, it suffices to jointly construct Ψ and $\widetilde{\Psi}(i, \omega)$ such that $\widetilde{\Psi}(i, \omega)$ is continuous i, and weakly increasing in both i and ω, and $\Psi(\omega) = \widetilde{\Psi}(\mu(\omega, \alpha), \omega)$ for all ω.

Since the construction of Ψ is a just a matter of definition given $\widetilde{\Psi}(i, \omega)$, the proof works toward constructing $\widetilde{\Psi}(i, \omega)$. We first construct it on the finite observed path. The construction is then extended to the remaining states and types. On the finite path, it is convenient to define monotone indices on ω and on i, respectively. Without loss of generality, we suppose that each on-path state ω_t is distinct. Otherwise, we let T denote the number of distinct observations of state variables and ignore repeated observations since they do not add new information. By reordering if necessary, we can define an index t with $t = 1, 2, ..., T$ such that

$\omega_t < \omega_{t+1}, \forall t < T$. The derived sequence of pivotal decision makers is defined as $\{i_t\}_{t=1}^{T}$ such that $i_t = \mu(\omega_t, \alpha)$. For the convenience of extending finite data to the whole range of states and types, we specify two fictional end-point observations as $(\omega_0, i_0, a_0) = (\min\Omega - 1, 0, \min A)$ and $(\omega_{T+1}, i_{T+1}, a_{T+1}) = (\max\Omega + 1, 1, \max A)$.[18]

Similarly, let N be the number of distinct elements in $\{i_t\}_{t=0}^{T+1}$ with $2 \le N \le (T+2)$. Define a second index n with $n = 1, 2, ..., N$ and the corresponding type sequence $\left\{\widetilde{i}_n\right\}_{n=1}^{N}$ with $\widetilde{i}_n \in \{i_t\}_{t=0}^{T+1}$ such that $\widetilde{i}_n < \widetilde{i}_{n+1}, \forall n < N$. In other words, n is a reordering of distinct elements in $\{i_t\}_{t=0}^{T+1}$ such that $\widetilde{i}_n$ is an increasing sequence. Notice that $\widetilde{i}_1 = 0$ and $\widetilde{i}_N = 1$.

We will construct $N \cdot (T+2)$ points of $\widetilde{\Psi}(i, \omega)$, all collectively denoted by $\{\widetilde{a}_{n,t}\}_{n=1,t=0}^{N,T+1}$, such that $\widetilde{a}_{n,t} = \widetilde{\Psi}\left(\widetilde{i}_n, \omega_t\right)$.

Notice first that equilibrium requires that $\widetilde{a}_{n,t} = a_t$ if $\widetilde{i}_n = i_t$. This leaves $(N-1) \cdot (T+2)$ points free for construction. To complete the finite construction, we specify an explicit algorithm to construct a weakly increasing sequence $\{\widetilde{a}_{n,t}\}_{n=1,t=0}^{N,T+1}$.

Algorithm 1 *A recursive algorithm to construct a weakly increasing* $\{\widetilde{a}_{n,t}\}_{n=1,t=0}^{N,T+1}$.

Step 0: Define an initial condition for $t = 0$ *as* $\widetilde{a}_{n,0} = a_0 = \min A$, $\forall 1 \le n \le N$.

Step 1: For observation t *with* $1 \le t \le T$, *find* $1 \le n_t^* \le N$ *such that* $\widetilde{i}_{n_t^*} = i_t$. *Let* $\widetilde{a}_{n_t^*,t} = a_t$. *For* $1 \le n \le N$ *and* $n \ne n_t^*$, *define* $\widetilde{a}_{n,t}$ *as an average of two points*

$$\widetilde{a}_{n,t} = \frac{1}{2}\left(\widetilde{a}_{n,t}^{\min} + \widetilde{a}_{n,t}^{\max}\right),$$

where $\widetilde{a}_{n,t}^{\min}$ *and* $\widetilde{a}_{n,t}^{\max}$ *are defined from* $\left\{\widetilde{a}_{n_t^*,t}, \{\widetilde{a}_{n,t-1}\}_{n=1}^{N}\right\}$ *in a recursion starting from* n_t^* *as*

$$\widetilde{a}_{n,t}^{\min} = \begin{cases} \max\{\widetilde{a}_{n-1,t}, \widetilde{a}_{n,t-1}\} & \text{if } n > n_t^*, \\ \widetilde{a}_{n,t-1} & \text{if } n < n_t^* \end{cases}$$

and

$$\widetilde{a}_{n,t}^{\max} = \left\{ \begin{array}{ll} \min\limits_{\left\{t' : T+1 \ge t' > t, i_{t'} \ge \widetilde{i}_n\right\}} \{a_{t'}\} & \text{if } n > n_t^*, \\ \min\left\{\widetilde{a}_{n+1,t}, \min\limits_{\left\{t' : T+1 \ge t' > t, i_{t'} \ge \widetilde{i}_n\right\}} \{a_{t'}\}\right\} & \text{if } n < n_t^*. \end{array} \right\}.$$

[18]The specific values of ω_0 and ω_{T+1} are not essential as long as they satisfy $\omega_0 < \min\Omega$ and $\omega_{T+1} > \max\Omega$.

Step 2: If $t < T$, let $t = t + 1$ and go back to Step 1; else go to Step 3.
Step 3: Let $\widetilde{a}_{n,T+1} = a_{T+1} = \max A$, $\forall 1 \leq n \leq N$ and stop.

For each $1 \leq t \leq T$, the Algorithm starts by producing the realized equilibrium policy outcome, $\widetilde{a}_{n_t^*,t} = a_t$. Starting from n_t^*, the Algorithm then proceeds to a two-way recursion towards both the left and right sides of n_t^*. It is easy to see that the Algorithm produces a non-empty sequence of real numbers. In addition, $\widetilde{a}_{n,t} \in A$, $\forall n, t$, since every operation involved, including min, max and mean, is a closed operation. Notice that $\widetilde{a}_{n,0} = \min A$ and $\widetilde{a}_{n,T+1} = \max A$. As a result, we only need to check that $\{\widetilde{a}_{n,t}\}_{n=1,t=1}^{N,T}$ is a weakly increasing sequence in (n, t).

Start from $t = 1$ and we prove the weak monotonicity of $\widetilde{a}_{n,t}$ in n in two steps.

Step 1: $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t} \geq \widetilde{a}_{n,t}^{\min}$ for every $n \neq n_t^*$. Because $\widetilde{a}_{n,t} = \frac{1}{2}\left(\widetilde{a}_{n,t}^{\min} + \widetilde{a}_{n,t}^{\max}\right)$, it suffices to show that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t}^{\min}$. We prove the fact for several cases of n. First, $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t}^{\min}$ for $1 \leq n < n_t^*$. From Step 0 of the Algorithm, it follows that $\widetilde{a}_{n,t}^{\max} \geq \min A = \widetilde{a}_{n,t-1} = \widetilde{a}_{n,t}^{\min}$ for every $1 \leq n < n_t^*$. Second, $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t}^{\min}$ for $n = n_t^* + 1$. Recall that $a_{t'} \geq a_t$ whenever $t' > t$ and $i_{t'} \geq i_t$. Hence, by taking the minimum we have $\widetilde{a}_{n,t}^{\max} \geq a_t$. In addition, $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t-1} = \min A$. For $n = n_t^* + 1$, it follows that $\widetilde{a}_{n,t}^{\max} \geq \max\{a_t, \widetilde{a}_{n,t-1}\} = \max\{\widetilde{a}_{n-1,t}, \widetilde{a}_{n,t-1}\} = \widetilde{a}_{n,t}^{\min}$. Third, $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t}^{\min}$ for $n_t^* + 1 < n \leq N$. For $n = n_t^* + 2$, notice that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n-1,t}^{\max} \geq \widetilde{a}_{n-1,t}$, where the first inequality follows because $\widetilde{i}_{n+1} > \widetilde{i}_n$ so that the set for min operation in the former is a subset of the latter, and the second inequality from the last result $\widetilde{a}_{n-1,t}^{\max} \geq \widetilde{a}_{n-1,t}^{\min}$ so that $\widetilde{a}_{n-1,t}^{\max} \geq \widetilde{a}_{n-1,t}$ for $n - 1 = n_t^* + 1$. Using this and the fact that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t-1}$, the same argument as in the previous step can establish that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t}^{\min}$ for $n = n_t^* + 2$. By induction, the same inequality holds for every $n_t^* + 1 < n \leq N$.

Step 2: $\widetilde{a}_{n,t}$ is weakly increasing in n for $t = 1$. From the construction, $\widetilde{a}_{n,t}^{\min} \geq \widetilde{a}_{n-1,t}$ for $n > n_t^*$ and $\widetilde{a}_{n,t}^{\max} \leq \widetilde{a}_{n+1,t}$ for $n < n_t^*$. Since $\widetilde{a}_{n,t}^{\min} \leq \widetilde{a}_{n,t} \leq \widetilde{a}_{n,t}^{\max}$ as shown in Step 1, we have $\widetilde{a}_{n-1,t} \leq \widetilde{a}_{n,t} \leq \widetilde{a}_{n+1,t}$.

For $1 < t \leq T$, the weak monotonicity of $\widetilde{a}_{n,t}$ in n is shown from an induction argument. Specifically, for each $t > 1$, we assume that $\widetilde{a}_{n,t-1}^{\max} \geq \widetilde{a}_{n,t-1} \geq \widetilde{a}_{n,t-1}^{\min}$ for every $n \neq n_{t-1}^*$, and $\widetilde{a}_{n,t-1}$ is weakly increasing in n, as derived for $t = 1$. Then we revisit the proof of Step 1 and Step 2 as in $t = 1$. It is easy to see that Step 2 is intact, provided that Step 1 holds. For Step 1, a close reading of the proof for $t = 1$ reveals that we only need to reestablish that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t-1}$, which follows from a series of claims.[19]

Claim 1: $\min\limits_{\left\{t':t'>t-1, i_{t'} \geq \widetilde{i}_n\right\}} \{a_{t'}\} \geq \widetilde{a}_{n,t-1}$ for every $1 \leq n \leq N$ and $1 < t \leq T$. For $n \neq n_{t-1}^*$, $\min\limits_{\left\{t':t'>t-1, i_{t'} \geq \widetilde{i}_n\right\}} \{a_{t'}\} \geq \widetilde{a}_{n,t-1}^{\max} \geq \widetilde{a}_{n,t-1}$, where the first inequality holds by construction, and the second inequality is true from the assumption of induction. For $n = n_{t-1}^*$, recall that $a_{t'} \geq a_{t-1}$

[19]Recall that $\widetilde{a}_{n,t}^{\max} \geq \widetilde{a}_{n,t-1}$ holds trivially for $t = 1$ from Step 0 of the Algorithm, which cannot be taken as given any more for $1 < t \leq T$.

whenever $t' > t-1$ and $i_{t'} \geq i_{t-1} = \widetilde{i}_n$. Take the minimum to get $\min\limits_{\{t':t'>t-1, i_{t'} \geq \widetilde{i}_n\}} \{a_{t'}\} \geq a_{t-1} = \widetilde{a}_{n^*_{t-1},t-1} = \widetilde{a}_{n,t-1}$.

Claim 2: $\min\limits_{\{t':t'>t, i_{t'} \geq \widetilde{i}_n\}} \{a_{t'}\} \geq \widetilde{a}_{n,t-1}$ for every $1 \leq n \leq N$ and $1 < t \leq T$. Notice that $\min\limits_{\{t':t'>t, i_{t'} \geq \widetilde{i}_n\}} \{a_{t'}\} \geq \min\limits_{\{t':t'>t-1, i_{t'} \geq \widetilde{i}_n\}} \{a_{t'}\}$, since the set for min operation in the former is a subset of the latter. The result then follows from the Claim 1.

Claim 3: $\widetilde{a}_{n^*_t,t} \geq \widetilde{a}_{n^*_t,t-1}$ for every $t > 1$. Notice that $a_t \geq \min\limits_{\{t':t'>t-1, i_{t'} \geq \widetilde{i}_n\}} \{a_{t'}\}$ for $n = n^*_t$, since $a_{t'}$ with $t' = t$ and $i_t = \widetilde{i}_{n^*_t}$ is one member of the constraint set. From the Claim 1, we have $\widetilde{a}_{n^*_t,t} = a_t \geq \widetilde{a}_{n^*_t,t-1}$.

Claim 4: $\widetilde{a}^{\max}_{n,t} \geq \widetilde{a}_{n,t-1}$ for $1 \leq n < n^*_t$. From the definition of $\widetilde{a}^{\max}_{n,t}$ for $1 \leq n < n^*_t$ and Claim 2, we only need to prove that $\widetilde{a}_{n+1,t} \geq \widetilde{a}_{n,t-1}$. Furthermore, it suffices to show that $\widetilde{a}_{n+1,t} \geq \widetilde{a}_{n+1,t-1}$, because $\widetilde{a}_{n+1,t-1} \geq \widetilde{a}_{n,t-1}$ from the weak monotonicity assumption of induction for $t-1$. For $n = n^*_t - 1$, $\widetilde{a}_{n+1,t} = \widetilde{a}_{n^*_t,t} \geq \widetilde{a}_{n+1,t-1}$ from Claim 3. In addition, by repeating the Step 1 as in $t = 1$, we have $\widetilde{a}_{n,t} \geq \widetilde{a}^{\min}_{n,t} = \widetilde{a}_{n,t-1}$ for $n = n^*_t - 1$. But this implies $\widetilde{a}^{\max}_{n,t} \geq \widetilde{a}_{n,t-1}$ for $n = n^*_t - 2$. By induction, the result holds for any $n < n^*_t$.

Claim 5: $\widetilde{a}^{\max}_{n,t} \geq \widetilde{a}_{n,t-1}$ for $n^*_t < n \leq N$. This follows immediately from Claim 2.

To summarize, we just proved that the Algorithm produces a weakly increasing sequence in n for each $0 \leq t \leq T+1$. It remains to show that $\widetilde{a}_{n,t}$ is weakly increasing in t for every $1 \leq n \leq N$. From the construction of $\widetilde{a}^{\min}_{n,t}$, for any t and any $n \neq n^*_t$, $\widetilde{a}_{n,t} \geq \widetilde{a}^{\min}_{n,t} \geq \widetilde{a}_{n,t-1}$. For $n = n^*_t$, from the Claim 3, $\widetilde{a}_{n^*_t,t} \geq \widetilde{a}_{n^*_t,t-1}$. Consequently, $\widetilde{a}_{n,t} \geq \widetilde{a}_{n,t-1}$, $\forall t, n$. This finishes the verification of the Algorithm.

Having constructed the points $\{\widetilde{a}_{n,t}\}_{n=1,t=0}^{N,T+1}$ and corresponding regular grid points $(\{\widetilde{i}_n\}_{n=1}^N, \{\omega_t\}_{t=0}^{T+1})$ from the algorithm, all that remains is to extend the construction to the full function $\widetilde{\Psi}(i,\omega)$. For this purpose, a standard bilinear interpolating spline can be used (for an introduction to splines, see Judd (1998)). Specifically, for each $i \in [\widetilde{i}_n, \widetilde{i}_{n+1}]$ and $\omega \in [\omega_t, \omega_{t+1}]$, a unique bilinear piece can be constructed as

$$\widetilde{\Psi}(i,\omega) = b^0_{n,t} + b^1_{n,t}\, i + b^2_{n,t}\, \omega + b^3_{n,t}\, i\omega, \tag{15}$$

such that $\widetilde{\Psi}(\widetilde{i}_n, \omega_t) = \widetilde{a}_{n,t}$, $\widetilde{\Psi}(\widetilde{i}_n, \omega_{t+1}) = \widetilde{a}_{n,t+1}$, $\widetilde{\Psi}(\widetilde{i}_{n+1}, \omega_t) = \widetilde{a}_{n+1,t}$, and $\widetilde{\Psi}(\widetilde{i}_{n+1}, \omega_{t+1}) = \widetilde{a}_{n+1,t+1}$.

By construction, $\widetilde{\Psi}(i,\omega)$ is continuous in (i,ω). In addition, a bilinear spline preserves the monotonicity property in each dimension: if $\{\widetilde{a}_{n,t}\}_{n=1,t=0}^{N,T+1}$ is a weakly increasing sequence in n (resp. t), then the constructed $\widetilde{\Psi}(i,\omega)$ is weakly increasing in i for each fixed $\omega \in \Omega$ (resp. in ω for each fixed $i \in [0,1]$). Because of the symmetry in (i,ω), it suffices to show the property for i, or $\frac{\partial\widetilde{\Psi}(i,\omega)}{\partial i} = b^1_{n,t} + b^3_{n,t}\omega \geq 0$. Notice that $\widetilde{\Psi}(i,\omega)$ is linear in i for each fixed ω, in particular for each ω_t and ω_{t+1}. Hence, $\widetilde{a}_{n+1,t} \geq \widetilde{a}_{n,t}$ and $\widetilde{a}_{n+1,t+1} \geq \widetilde{a}_{n,t+1}$ imply that $\frac{\partial\widetilde{\Psi}(i,\omega_t)}{\partial i} = b^1_{n,t} + b^3_{n,t}\omega_t \geq 0$ and $\frac{\partial\widetilde{\Psi}(i,\omega_{t+1})}{\partial i} = b^1_{n,t} + b^3_{n,t}\omega_{t+1} \geq 0$. It immediately follows that $\frac{\partial\widetilde{\Psi}(i,\omega)}{\partial i} = b^1_{n,t} + b^3_{n,t}\omega \geq 0$ for every $\omega \in [\omega_t, \omega_{t+1}]$.

With the extension to the full function $\widetilde{\Psi}$, the proof of Theorem 2 is complete. ∎

Proof of Theorem 3. **Part 1**. We first suppose that Part 2 holds, and proceed to prove existence of a rationalizing α for any observed $\{a_t\}_{t=1}^T$ and $\{p_t, q_t\}$. Given the data, we can pick a sequence of pivotal decision maker $\{i_t\}_{t=1}^T$ such that inequality (11) holds for $i_t = \mu(\alpha, \omega_t)$. From the monotonicity and full range assumption on α in (B3), there exists a sequence of $\{\alpha_t\}_{t=1}^T$ with $\mu(\omega_t, \alpha_t) = i_t$. As a result, any $\alpha(\omega)$ with $\alpha(\omega_t) = \alpha_t$ rationalizes the data.

Part 2. Necessity was proved in the paper. To prove sufficiency, fix any data and any weighting function α that satisfy the inequalities in (11). Consider any policy rule $\Psi(\omega)$ consistent with the data and an admissible U of the form in (6) given in the Proof of Theorem 1. We now add the requirement that $\widetilde{\Psi}$ in (6) be increasing in i. This implies strict single crossing of U in $(a; i)$. To prove consistency of U with polling data comparing a_t to $\bar{a}$, it suffices to assume that the type p_t is indifferent between a_t and $\overline{a}$, i.e., $U(p_t, \omega_t, \overline{a};\ \Psi) = U(p_t, \omega_t, \Psi(\omega_t); \Psi)$. With some algebra, it reduces to

$$\widetilde{\Psi}(p_t, \omega_t) = \frac{1}{2}(\overline{a} + a_t),$$

where $\widetilde{\Psi}$ is the function associated with payoff function U as specified in (6). Similarly, it suffices to establish consistency of U with polling for a_t against $\underline{a}$ for each t by assuming that the type $1-q_t$ is indifferent between a_t and $\underline{a}$, i.e., $U(1-q_t, \omega_t, \underline{a};\ \Psi) = U(1-q_t, \omega_t, \Psi(\omega_t); \Psi)$, which reduces to

$$\widetilde{\Psi}(1 - q_t, \omega_t) = \frac{1}{2}(\underline{a} + a_t).$$

To prove that α rationalizes the data with policy rule Ψ under admissible payoff function U of the form in (6), it therefore suffices to construct function $\widetilde{\Psi}$ that satisfies:

(i) $\widetilde{\Psi}$ is increasing in i, and

(ii) $\widetilde{\Psi}$ satisfies the equation systems

$$\begin{pmatrix} \widetilde{\Psi}(1 - q_t, \omega_t) \\ \widetilde{\Psi}(\mu(\omega_t, \alpha), \omega_t) \\ \widetilde{\Psi}(p_t, \omega_t) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\underline{a} + a_t) \\ a_t \\ \frac{1}{2}(\overline{a} + a_t) \end{pmatrix} \qquad t = 1, \ldots, T. \tag{16}$$

Hence, fix any on-path ω_t. Then $\widetilde{\Psi}(i,\omega_t)$ can be found as a linear spline passing through five data points $(0,\widetilde{\Psi}(0,\omega_t))$,$(1-q_t,\frac{1}{2}(\underline{a}+a_t))$,$(\mu(\omega_t,\alpha),a_t)$, $(p_t,\frac{1}{2}(\overline{a}+a_t))$ and $(1,\widetilde{\Psi}(1,\omega_t))$, where

$$\widetilde{\Psi}(0,\omega_t)=\begin{cases}\frac{1}{2}(\underline{a}+a_t) & \text{if } 1-q_t=0\\ \\ \underline{a} & \text{if } 1-q_t>0\end{cases},$$

and

$$\widetilde{\Psi}(1,\omega_t)=\begin{cases}\overline{a} & \text{if } p_t<1\\ \\ \frac{1}{2}(\overline{a}+a_t) & \text{if } p_t=1\end{cases},$$

Notice that under the assumption $1-q_t<\mu(\omega,\alpha)<p_t$, the data points along i dimension are increasing since $\underline{a}<a_t<\overline{a}$. As a result, $\widetilde{\Psi}(i,\omega_t)$ is increasing in i for each ω_t. Since $\widetilde{\Psi}(i,\omega)$ is not restricted off-path, any $\widetilde{\Psi}(i,\omega)$ increasing in i will serve the purpose. For instance, the construction used in Theorem 1, $\widetilde{\Psi}(i,\omega_t)=i-\mu(\omega,\alpha)+\Psi(\omega)$, will work. This concludes the proof. ∎

Sufficiency Proof of Theorem 7. Let the construction of $\widetilde{\Psi}$ be the same as in Theorem 3. A careful reading of the proof in Theorem 3 reveals that the restrictions of polling data are similar, except for an increased number of data points that lie on the constructed linear spline. Because it is a straightforward adaptation of the existing argument, we omit the details and leave this as an exercise to the reader. ∎

References

[1] Acemoglu, D. and J. Robinson (2008), "Persistence of Powers, Elites and Institutions," *American Economic Review*, 98: 267-293.

[2] Afriat, S. (1967), "The Construction of a Utility Function from Expenditure Data," *International Economic Review*, 8: 67-77.

[3] Austen-Smith, D. (1987), "Interest Groups, Campaign Contributions, and Probabilistic Voting", *Public Choice* 54: 123-139.

[4] Bartels, L. (2008), *Unequal Democracy: The Political Economy of the New Gilded Age*, Princeton, NJ: Princeton University Press.

[5] Benabou, R. (2000), "Unequal Societies: Income Distribution and the Social Contract,"*American Economic Review*, 90(1): 96-129.

[6] Boldrin, M. and L. Montruchio (1986), "On the Indeterminacy of Capital Accumulation Paths," *Journal of Economic Theory*, 40: 26-39.

[7] Bourguignon, F., and T. Verdier (2000), "Oligarchy, Democracy, Inequality and Growth," *Journal of Development Economics*, 62: 287-313.

[8] Brown, D. and F. Kubler (2008), *Computational Aspects of General Equilibrium Theory: Refutable Theories of Value*, Springer-Verlag.

[9] Campante, F. (2008), "Redistribution in a Model of Voting and Campaign Contributions," mimeo, Harvard University.

[10] Chiappori, P.-A. and J.-C. Rochet (1987), "Revealed Preference and Differentiable Demand, *Econometrica*, 55: 687-91.

[11] Coate, S. (2004), "Political Competition with Campaign Contributions and Informative Advertising", *Journal of the European Economic Association*, 2: 772-804.

[12] Debreu, G. (1974). "Excess Demand Functions," *Journal of Mathematical Economics*, 1: 15-21.

[13] Degan, A. and A. Merlo (2009), " Do Voters Vote Ideologically?" *Journal of Economic Theory*, 144:1868-94.

[14] Gans, J. and M. Smart (1996), "Majority voting with single-crossing preferences," *Journal of Public Economics*, 59: 219-237.

[15] Grossman, G. and E. Helpman (1996), "Electoral Competition and Special Interest Policies", *Review of Economic Studies*, 63: 265-286.

[16] Judd, K. (1998), *Numerical Methods in Economics*, Cambridge: MIT Press.

[17] Kalandrakis, T. (2010), "Rationalizable Voting," *Theoretical Economics*, 5:93-125.

[18] Mantel, R. (1974), "On the characterization of aggregate excess demand," *Journal of Economic Theory* 7: 348—353.

[19] Prat, A. (2002), "Campaign Spending with Office-Seeking Politicians, Rational Voters, and Multiple Lobbies", *Journal of Economic Theory*, 103: 162-189.

[20] Richter, M. K. (1966), "Revealed Preference Theory," *Econometrica*, 34: 635—645.

[21] Rosenstone, S. and J.M. Hansen (1993), *Mobilization, Participation and Democracy in America*, Macmillan, New York.

[22] Rothstein, P. (1990), "Order Restricted Preferences and Majority Rule," *Social Choice and Welfare*, 7: 331-42.

[23] Sonnenschein, H. (1973), "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?", *Journal of Economic Theory* 6: 345—354.

[24] Varian, H. (1982), "The Non-Parametric Approach to Demand Analysis," *Econometrica*, 50: 945-974.

[25] Varian, H. (2006), "Revealed Preference," In: Michael Szenberg, Lall Ramrattan, Aron A. Gottesman (Eds.), *Samuelsonian Economics and the Twenty-First Century*, Oxford University Press, pp. 99-115.