Inequality Aversion and Preferences for Redistribution: New Tests of Political Economy Theories

Ben Goodrich, Columbia

September 16, 2011

Abstract

To what extent do preferences for redistribution reflect the private self-interest of the voter? Many political economy models assume that once private self-interest is properly accounted for, the residual variation in preferences for redistribution is negligible. Many theories outside of the traditional political economy literature assume that private self-interest is one among several motivations to support or oppose redistribution.

The primary goal of this paper is to provide compelling empirical evidence for the latter view. To do so, it simultaneously models preferences for domestic redistribution, progressive taxation, and international redistribution, using a newly-developed quantitative method. If preferences for redistribution were merely private self-interest, then no one in a wealthy country would support international redistribution. In fact, there is considerable within-country variation in preferences for international redistribution, and the new quantitative method demonstrates that preferences for international redistribution covary with preferences for domestic redistribution and progressive taxation, even after controlling for private self-interest. Moreover, the within-country variation in preferences for domestic redistribution that is attributable to private self-interest is small relative to the total amount of systematic, within-country variation in preferences for domestic redistribution across more than twenty advanced industrial and Eastern Europen countries.

This empirical finding should encourage political economists to rethink the standard theoretical model(s) for policy preferences. In short, there is more to be gained from a better understanding of the variance in preferences attributable to non-economic motivations, a topic where this paper only scratches the surface. Experiments in neuroscience, economics, psychology are begining to provide a clearer picture of these non-economic motivations, which should help political economists integrate these insights into a theoretical model that makes testable predictions about policy in equilibrium.

1 Introduction

One of the classic debates both within political science and across the social sciences concerns the extent to which the political behavior of individuals is driven by private self-interest. This debate responded to the fundamental question, "If the majority of the population in a democracy is poor, why do the poor not vote to equalize income after taxes and transfers?" Answers to this question differ, depending on one's view of human nature.

Economic theory traditionally argues that private self-interest is the primary if not exclusive motivation for all individual behavior. More formally, the utility function for voter j depends only on constants and variables are indexed by j. The usual starting point for this literature is the model in Meltzer and Richard (1981), which proved that self-interested poor voters would not prefer the maximum tax *rate*, but rather the tax rate that maximized tax *revenue*. Thus, the Meltzer-Richard model shows that it is not in the self-interest of the poor to fully equalize after-tax income because there would be no tax revenue to redistribute.

The alternative theory denies the exclusivity of traditional political economy and argues that the utility function for voter j can also depend on the outcomes of other voters. The alternative literature also has a long history in many fields, but the modern touchpoint is Fehr and Schmidt (1999), which popularized the term "inequality aversion" to capture the notion that some people dislike societal inequality in much the same way that some people are thought to have a distaste for risk. Under this view, equity is a "good", presumably with an upward-sloping supply schedule that, along with the demand curve, determines the equilibrium level of redistribution.

Several authors, such as Fong (2001), Lübker (2007), Georgiadis and Manning (2007), and Alesina and Giuliano (2009) have found that survey data do not support the traditional political economy model at the individual level. To extend upon their empirical results, we consider preferences for both within-country redistribution and cross-country redistribution. Our main empirical result is that the motivation for within-country redistribution is qualitatively similar to the motivation for cross-country redistribution, which is difficult to reconcile with traditional political economy models because cross-country redistribution is not in the donor's private self-interest. Not only is inequality aversion relevant, in many advanced industrial countries, a plurality of the variation in support for domestic redistribution is attributable to variation in inequality aversion. This finding implies that there is much to be gained by further investigation into the

nature of inequality aversion and why some people are more inequality averse than others.

More generally, the empirical success of the alternate view implies that we should think of redistribution, in no small part, as a *social* issue as well as a personal economic issue (see, for example, Gilens 2000). There are many polarizing social issues — such as abortion, gay marriage, etc. — where voters have strong preferences over policies, even if they are unlikely to be affected by these policies on a direct, personal level. Although almost everyone pays some taxes and receives some benefits from the government and thus has a personal financial stake in welfare-state policy, these policies also have major social ramifications that are relevant to the preferences of voters that place some weight on the social dimension.

These results also have implications for several related literatures. First, the empirical results provide some external validation of the experimental outcomes in distribution games. While the survey data do not speak directly to political *behavior*, the survey data suggest that the behavior in lab experiments is relevant for politics. If so, this finding implies that we need a different approach to modeling aggregate equilibria than the simple, single-crossing mechanism that are prevalent in the political economy literature. Second, the results support the main conclusion in Ansolabehere, Rodden and Snyder (2008) that survey respondents express reasonably consistent preferences across survey items, but this consistency is only revealed when several survey items are modeled simultaneously. However, we avoid some methodological weaknesses in Ansolabehere, Rodden and Snyder (2008) by utilizing an algorithm developed in Goodrich (2009*a*), and the results in this paper suggest that the algorithm is useful in practical research situations. Thus, it is plausible that a similar results would hold for other welfare-state policies that have major distributive implications, such as pensions or health insurance.

The outline of this paper is straightforward. First, we discuss the two theoretical approaches to preferences for redistribution. The remaining sections build to the main empirical result that inequality aversion is an important concept in explaining preferences for both international and domestic redistribution. The conclusion considers the broader implications for the political economy literature.

2 Modeling Two Competing Theories of Preferences for Redistribution

This section contrasts "traditional" political economy theories with "alternative" inequality aversion theories. One important contrast is that political economy traditionally assumes that a voter's utility function for

a policy depends only on constants and private variables, like the voter's *own* consumption, *own* leisure, etc. The alternative approach does not necessarily deny that self-interest is relevant but assumes that a voter's utility also depends on the payoffs of *others*. Of course, it is quite possible for political economists to derive the implications of utility functions that depend on payoffs of others, such as in Corneo and Grüner (2000).

The second distinguishing characteristic is that political economy traditionally constructs its models in such a way as to ensure that there is a Condorcet-winning policy in equilibrium. Often the equilibrium is a consequence of the single-crossing theorem (see Gans and Smart 1996), which says that there will be a Condorcet-winning policy under a majoritarian voting rule if policy preferences can be ordered by a single-characteristic of the agents, such as their own income. The alternative approach tends to use evolutionary or otherwise computational models to describe aggregate behavior if that is of interest.

The two senses of "traditional" are complementary in that a Condorcet-winning policy will typically exist only if a voter's utility depends only on a very small number of private variables. In this paper, we are primarily concerned with challenging tradition in the first sense, but the results have implications for how political economists can profitably study aggregate effects that are discussed in the concluding section.

The best example of a traditional political economy model is the model in Meltzer and Richard (1981). It is "traditional" in both senses. First, utility depends on own consumption and own leisure with the constraint that there are only twenty-four hours in the day, which much be allocated between labor and leisure. Second, in equilibrium, the decisive voter is the median income earner under majority rule with universal suffrage.

To motivate our empirical work, we derive a model that is similar to the Meltzer-Richard model now. Let a continuum of voters be defined on the $(0, 1)$ interval. Fiscal policy involves a proportional tax rate, $\tau \in [0, 1]$, and a lump-sum benefit, which is the mechanism used in the Meltzer-Richard model and many traditional political economy models. In contrast to the Meltzer-Richard model, we assume that hours worked are *fixed* by employers. Each individual simply has a binary choice to work or not. In other words, the labor supply effects of fiscal policy go through participation in the labor market rather than through marginal changes in the number of hours worked. This setup is arguably more realistic in light of the empirical evidence that suggests most workers have only a limited ability to choose their total number of hours that they work.

Thus, we define

$$\omega_j(\tau) = \begin{cases} 1 & \text{if } j \text{ is employed} \\ 0 & \text{if } j \text{ is unemployed,} \end{cases}$$

which is a function of τ, albeit not a differentiable function of τ. Working also entails some fixed cost $c \geq 0$, which could be interpreted as the opportunity cost of forgone leisure and does not vary across workers in these models. Let $\rho(\tau)$ be the "unemployment rate", i.e. the proportion of the population that does not work when the tax rate is τ (and implicitly when the cost of working is c). Conversely, $[1 - \rho(\tau)]$ is the "employment rate", the proportion of the population that works, given τ.

Let $Q(i)$ be a quantile function that maps from the $(0, 1)$ interval to the amount of money (Y) nature would allocate to i if i works. In other words, $Q(i)$ can be interpreted as nature's productivity function, i is the proportion of the population with less productivity, and the median productivity is $Q(0.5)$. A quantile function is the inverse of a cumulative distribution function (CDF), i.e. $F(Q(i)) = i$. Let $y_i = \omega_i(\tau)Q(i)$ be the *realized* wage of a worker. If $\omega_i(\tau) = 1$, then $Q(F(y_i)) = y_i$, and if $\omega_i(\tau) = 0$, then $Q(i) > y_i = 0$. Two useful identities are that $\int_0^1 Q(i)di \equiv \mathbb{E}(Y)$ and $\frac{\int_{\rho(\tau)}^1 Q(i)di}{1-\rho(\tau)} \equiv \mathbb{E}(Y | \omega(\tau) = 1)$. For compactness, we suppress the di henceforth but always integrate over i (voters), while using j as an alternate index for voters.

Here the utility function for voter j is composed of two parts: after-tax wages less the fixed cost of working and the lump-sum subsidy each person receives from the government's tax revenue:

$$\begin{aligned} U_j(\tau; Q(j), c) =& \omega_j(\tau) \left[\overbrace{[1-\tau]\, Q(j)}^{\text{after-tax wages}} \underbrace{-c}_{\text{cost}} \right] + \overbrace{\int_{\rho(\tau)}^1 \tau Q(i)}^{\text{subsidy}}, \\ =& \omega_j(\tau) \left[[1-\tau]\, Q(j) - c \right] + \tau \left[1 - \rho(\tau) \right] \mathbb{E}(Y | \omega(\tau) = 1). \end{aligned}$$

The second term is equal to the lump-sum subsidy that everyone receives under a balanced budget constraint. If $\rho(\tau) = 0$, which is to say that everyone works, then the subsidy simplifies to $\tau\mathbb{E}(Y)$. If $\rho(\tau) = 1$, which is to say that no one works, then there is no tax revenue and no subsidy.

However, $\rho(\tau)$ is endogenous because τ is the choice variable for voters. Under this utility function, $\omega_j(\tau) = 1$ iff $[1-\tau]\, Q(j) \geq c$. Hence, $\rho(\tau) \equiv Q^{-1}\left(\frac{c}{1-\tau}\right) \equiv F\left(\frac{c}{1-\tau}\right)$, and $\frac{\partial \rho(\tau)}{\partial \tau} = \frac{\partial F\left(\frac{c}{1-\tau}\right)}{\partial \frac{c}{1-\tau}} \times$

$\frac{\partial \frac{c}{1-\tau}}{\partial \tau} = \frac{cf\left(\frac{c}{1-\tau}\right)}{[1-\tau]^2} \geq 0$, where $f\left(\frac{c}{1-\tau}\right)$ is the PDF of nature's productivity distribution evaluated at the point of indifference to working. The (second) fundamental theorem of calculus along with the chain rule then implies that $\frac{\partial \int_{\rho(\tau)}^{1} Q(i)}{\partial \tau} = \frac{\partial \int_{\rho(\tau)}^{1} Q(i)}{\partial \rho(\tau)} \times \frac{\partial \rho(\tau)}{\partial \tau} = -Q\left(\rho(\tau)\right) \times \frac{\partial \rho(\tau)}{\partial \tau} = -Q\left(F\left(\frac{c}{1-\tau}\right)\right) \times \frac{\partial \rho(\tau)}{\partial \tau} = -\frac{c}{1-\tau} \times \frac{cf\left(\frac{c}{1-\tau}\right)}{[1-\tau]^2} = \frac{-c^2 f\left(\frac{c}{1-\tau}\right)}{[1-\tau]^3} \leq 0$.

Among those that work ($\omega_j(\tau) = 1$), the derivative of the utility function with respect to τ is

$$\frac{\partial U_j\left(\tau; Q(j), c\right)}{\partial \tau} = -y_j - \tau \frac{c^2 f\left(\frac{c}{1-\tau}\right)}{[1-\tau]^3} + [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right).$$

Setting this first derivative equal to zero and rearranging yields a first-order condition where workers choose τ to satisfy

$$y_j = [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right) - \frac{\tau c^2 f\left(\frac{c}{1-\tau}\right)}{[1-\tau]^3}.$$

Since all the terms are non-negative, this first-order condition can never satisfied among the "rich", who are defined as voters such that $y_j > [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right)$. The rich prefer the minimum feasible tax rate, which is $\tau = 0$. The "middle class" is then defined as workers for whom $y_j < [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right)$ and hence can choose $\tau \in (0, 1)$ to satisfy the first-order condition. The voter that defines the cutpoint between the middle class and the rich prefers $\tau = 0$, i.e. for that voter $y_j = [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right)$.

The "poor" are defined as non-workers ($\omega_j(\tau) = 0$) and have no wages. Thus, they choose τ to satisfy a special case of the first-order condition where $y_j = 0$:

$$0 = [1-\rho(\tau)]\,\mathbb{E}\left(Y \middle| \omega(\tau) = 1\right) - \tau \frac{c^2 f\left(\frac{c}{1-\tau}\right)}{[1-\tau]^3}.$$

The poor all prefer the same tax rate, but in general do not favor the maximum tax rate of unity due to the disemployment effects. The poor choose τ to maximize tax *revenue*, which requires equating the marginal gain from increasing the tax rate with the marginal cost of reducing the tax base.

It is easy to show that preferences for τ can be (weakly) ordered by wages or productivity. Hence, the single-crossing theorem tells us that there is a unique equilibrium under majority rule with universal suffrage where the decisive voter is the median wage earner. Among countries with pure majority rule and universal

suffrage, this simple model predicts that cross-country differences in redistribution depend on the shape of the productivity distribution (and on c), which is similar to the result in Meltzer and Richard (1981). The Meltzer-Richard model yields slightly sharper predictions, but the main result that preferences for τ can be ordered by wages is the same in both cases. In other words, this model is in the (large) family of political economy models in the Meltzer-Richard tradition that imply preferences can be ordered by wages.

Does the Meltzer-Richard model or a similar one fit the data well? At the aggregate level, it is widely recognized that redistribution is not necessarily positively associated with wage inequality, even in majoritarian countries. There are many refinements to the Meltzer-Richard model at the individual level that perhaps add variables to the model but continue to assume preferences are determined exclusively by private self-interest. However, these enhanced models do not predict an individual's preferences well. For example, Fong (2001) includes twenty-five explanatory variables and can explain only nine percent of the variation in preferences for redistribution in the United States. In some specifications Fong (2001) cannot reject the null hypothesis that private self-interest is completely irrelevant to preferences for redistribution. Similarly, Georgiadis and Manning (2007) models British preferences for redistribution with thirty six explanatory variables that are intended to capture private self-interest but can explain only ten percent of the variation, even with yearly dummy variables. Cusack, Iversen and Rehm (2006) is the most extensive test of the insurance hypothesis using data from many countries, and about five percent of the pooled variation in preferences is explained by the model, despite the inclusion of country and year dummies. Finally, Alesina and Giuliano (2009) models US and cross-country preferences for redistribution and finds that private self-interest explains about ten percent of the variance.

Thus, traditional political economists face a dilemma. Either individual preferences for redistribution (at least as measured in surveys) are almost entirely noise or preferences for redistribution are not exclusively or even primarily a function of theorized private variables (or at least those collected by surveys). In short, one can blame the survey and / or blame the theory, but in either case, the inability of the theoretical model to hold at the individual-level is sufficient to prevent the aggregate-level predictions from holding.

One difficulty with defining the "alternative" view is that it fuses the ideas of heterogeneous groups, such as psychologists, sociologists, neuroscientists, anthropologists, philosophers, political scientist that study public opinion, experimental economists, and probably others. Thus, it is difficult to succinctly characterize

the views of these researchers. They are united by the belief that researchers must take *social* considerations into account. In particular, whether an individual favors or opposes a policy depends not only on how that policy would affect his or her own income but also how that policy would affect others and whether these effects are legitimate. Of course, which group of people constitutes the relevant "others" can vary from person to person or from situation to situation.

In political science, perhaps the most explicit statement of the alternative view comes from Gilens (2000), which argues:

> Politics is often viewed, by elites at least, as a process centered on the question "who gets what." For ordinary Americans, however, politics is more often about "who deserves what," and the welfare-state is no exception. When they evaluate government policy, the foremost question in their minds is "what policy is best," not "what policy is best for me." Like much of politics, support or opposition to the various programs of the welfare state turns on issues of merit and deservedness ... (2)

In short, Gilens (2000) rejects the idea that preferences are driven exclusively by private self-interest, and presents some empirical evidence in support of the above hypothesis for the United States. In the empirical part of this paper, we attempt to extend Gilens' (2000) finding to advanced industrial countries generally.

To motivate our empirical strategy, we briefly consider lab experiments on distributional preferences. One of the simplest lab experiments is the Dictator Game (DG). In the DG, there are two players and player 1 is given a sum of money to divide between the two players. Player 1 makes a decision over how to divide the money and the two players keep the resulting payoffs. Traditional economic logic implies that player 1 will not give any money to player 2, but many player 1s offer player 2 a nontrivial amount of money.[1] In response, Fehr and Schmidt (1999) proposes the following utility function in attempt to characterize behavior across a variety of lab experiments,

$$U_j(\mathbf{z}) = z_i - \frac{\alpha_j}{n-1}\sum_{j\neq i}\max(z_i - z_j, 0) - \frac{\beta_j}{n-1}\sum_{j\neq i}\max(z_j - z_i, 0),$$

[1] Fehr and Schmidt (1999) summarizes previous results in the literature from the DG: twenty to forty percent of player 1s give no money to player 2, an approximately equal percentage gives half the money to player 2, and the rest gives a positive share but less than half to player 2. To account for this variation in behavior, researchers are essentially forced to assume that player 1's utility also depends on player 2's payoffs. There are variations on the DG with third parties that can engage in costly punishment, other distribution games, public goods games, etc., but traditional economic predictions tend to fare poorly in all of them.

where α_j and β_j respectively capture j's aversion to payoffs that disfavor and favor j relative to i and where $\mathbf{z}$ is a n-vector of payoffs in the game. Although some have criticized the details of this utility function, it seems to be roughly consistent with observed behavior in several different kinds of lab experiments, implying the distribution of α_j and β_j can be (and has been) calibrated to the results of an experiment.

To model a continuum of individuals in a welfare state, we need to rewrite the Fehr-Schmidt utility function somewhat. The inequality aversion parameters, α_j and β_j, affect policy preferences but do not affect j's private decision to work. If j works — that is if $[1-\tau]\,Q(j) \geq c$ — the modified utility function is

$$\begin{aligned} U_j\left(\tau; Q(j), c\right) \quad = \quad & \left[\overbrace{[1-\tau]\,Q\left(j\right)}^{\text{after-tax wages}} \underbrace{-c}_{\text{cost}}\right] + \overbrace{\tau \int_{\rho(\tau)}^{1} Q(i)}^{\text{subsidy}} - \\ & \overbrace{\alpha_j \left[-\left[1-\tau\right] y_j + \left[1-\tau\right] \int_{j}^{1} Q(i)\right]}^{\text{disfavorable outcomes to } j} - \\ & \underbrace{\beta_j \left[\left[1-\tau\right] \int_{0}^{\rho(\tau)} y_j + \left[1-\tau\right] \int_{\rho(\tau)}^{j} \left[y_j - Q(i)\right]\right]}_{\text{favorable outcomes to } j}, \end{aligned}$$

which simplifies to

$$U_j\left(\tau; Q(j), c\right) \quad = \quad \left[\left[1-\tau\right] y_j - c\right] + \tau \int_{\rho(\tau)}^{1} Q(i) - \alpha_j \left[1-\tau\right] \left[-y_j + \int_{j}^{1} Q(i)\right] - \beta_j \left[1-\tau\right] \left[j y_j - \int_{\rho(\tau)}^{j} Q(i)\right]$$

Among those that work, the derivative with respect to τ is

$$\begin{aligned} \frac{\partial U_j\left(\tau; Q(j), c\right)}{\partial \tau} \quad = \quad & -y_j - \tau \frac{c^2 f\left(\frac{c}{1-\tau}\right)}{\left[1-\tau\right]^3} + \int_{\rho(\tau)}^{1} Q(i) + \\ & \alpha_j \left[-y_j + \int_{j}^{1} Q(i)\right] + \\ & \beta_j \left[j y_j - \int_{\rho(\tau)}^{j} Q(i)\right] - \\ & \beta_j \left[1-\tau\right] \frac{c^2 f\left(\frac{c}{1-\tau}\right)}{\left[1-\tau\right]^3}, \end{aligned}$$

which, when set equal to zero implies the first-order condition

$$\begin{aligned} y_j\left(1+\alpha_j-j\beta_j\right) &= \frac{c^2 f\left(\frac{c}{1-\tau}\right)}{\left[1-\tau\right]^3}\left(\beta_j\left[1-\tau\right]-\tau\right)+\left[1+\alpha_j\right]\left[1-\rho(\tau)\right]\mathbb{E}\left(Y|\,\omega(\tau)=1\right)+ \\ & \alpha_j\int_j^1 Q(i)-\beta_j\int_{\rho(\tau)}^j Q(i). \end{aligned}$$

All non-workers have the same post-fisc income, so $\max\left(z_j-z_i,0\right)=0$ for all unemployed j. In other words, their utility function is simply

$$\begin{aligned} U_j\left(\tau\right) &= \tau\int_{\rho(\tau)}^1 Q(i)-\alpha_j\left[1-\tau\right]\int_{\rho(\tau)}^1 Q(i), \\ &= \left[\tau-\alpha_j\left[1-\tau\right]\right]\int_{\rho(\tau)}^1 Q(i). \end{aligned}$$

Thus, the first-order condition is

$$\left[\left[1+\alpha_j\right]\tau-\alpha_j\right]\frac{c^2 f\left(\frac{c}{1-\tau}\right)}{\left[1-\tau\right]^3} = \left[1+\alpha_j\right]\left[1-\rho(\tau)\right]\mathbb{E}\left(Y|\,\omega(\tau)=1\right).$$

Among workers, if $\alpha_j=0=\beta_j$, then their first-order condition simplifies to that in our traditional political economy model. If $\alpha_j,\beta_j>0$, then the τ preferred by a working j is larger in the inequality aversion model. If $\alpha_j,\beta_j<0$, then the τ preferred by j is smaller in this model than in the traditional political economy model. In other words, we allow for "negative inequality aversion" in the same way that "negative risk aversion" implies risk-seeking behavior. Similarly, for non-workers, if $\alpha_j=0$, then this first-order condition simplifies to that in our traditional political economy model. If $\alpha_j>0$, the τ preferred by non-workers is larger and conversely if $\alpha_j<0$. Thus, our empirical strategy is essentially to test the null hypothesis from the traditional political economy literature that inequality aversion is either constant or irrelevant to preferences for redistribution.

The result for non-workers raises an important point that is rarely mentioned in the literature. In the Fehr-Schmidt model, it is quite possible for the poor to support redistribution both out of private self-interest *and* inequality aversion. In surveys, when the poor express support for redistribution, it is almost always interpreted as evidence in favor of a traditional political economy model. This perhaps represents the

simplest explanation of the behavior. However, it is possible that those same poor people would continue to support (some) redistribution even if they were rich, due to the inequality aversion parameters. Conversely, it is possible that rich people who oppose redistribution would continue to oppose redistribution (somewhat) even if they were poor. Thus, if the Fehr-Schmidt model is correct, the coefficient on income in a traditional political economy regression means little unless inequality aversion is somehow controlled for. Controlling for inequality aversion in survey data is difficult, which is why most of the recent literature has focused on randomized experiments, but we try to establish the external validity of these experiments with a unique cross-country survey.

The alternative model includes our traditional political economy model as a special case. However, unlike traditional political economy models, the single-crossing condition is not satisfied in the alternative model, except in unlikely special cases. The aggregate implications of the alternative model go beyond the scope of this paper, but suffice it to say that aggregate data is insufficient to test this model against traditional political economy models. Thus, in section 4, we test the two models at the individual-level.

3 Methods for Estimating Multiple Preferences

In order to be analytically useful, it is essential that the proposed utility function be appropriate for *several* different phenomena, thereby making it a general theory rather than a particularistic one. Hence, we have to also analyze *other* political preferences that are also thought to be a function of the same explanatory variables in order to judge between models in the Meltzer-Richard and Fehr-Schmidt families. However, considering preferences over multiple issues is problematic for the second sense of "traditional" political economy because it tends to preclude the existence of a Condorcet-winning bundle. Nevertheless, this inconvenience should be seen as an opportunity because simultaneously modeling multiple preferences provides additional leverage on the theoretical questions raised in this paper.

The Appendix closely follows a derivation in Fessler (1996) to show that factor analysis — which is an empirical tool generally associated with psychologists, sociologists, etc. — is perfectly appropriate for studying utility-maximizing agents, which is the hallmark of microeconomics. The traditional political economy camp and the alternative camp have fundamentally different priors about the number of factors needed to explain preferences for policies that affect the distribution of income. Political economists traditionally

do not utilize factor analysis, but their theory implies that the only factor that is relevant to preferences for redistribution is self-interest, which can be ascertained by where a voter falls in the productivity distribution. The alternative camp rejects the sufficiency assumption of the traditional political economy, and in particular, the Fehr-Schmidt model assumes three factors are necessary and sufficient. Thus, a great deal of leverage on the theoretical questions raised in this paper can be obtained simply by estimating the number of factors.

Ironically, productivity is essentially a *latent* variable that cannot be directly observed, so political economists traditionally assume that wages are a monotonic function of productivity in order to justify their regressions of preferences on wages. However, factor analysis is the canonical tool when researchers observe outcomes of variables that are thought to be functions of unobservable factors. In other words, in the context of the Meltzer-Richard model, we could model both wages and preferences for redistribution as a function of one latent factor, which is thought to be productivity, and this empirical strategy would be *more* consistent with the Meltzer-Richard theory than is a regression of preferences for redistribution on wages (plus whatever additional control variables that are not part of the Meltzer-Richard model the researcher decides to include).

The factor analysis model in the population can be written in matrix form as $\mathbf{\Sigma} = \mathbf{\Lambda \Upsilon \Lambda}' + \mathbf{\Theta}$, where $\mathbf{\Sigma}$ is a correlation matrix among observed preferences, $\mathbf{\Upsilon}$is a correlation matrix among latent factors, $\mathbf{\Lambda}$ is a matrix of linear regression coefficients of the latent factors on the observed variables, and $\mathbf{\Theta}$ is a diagonal covariance matrix among the errors in predicting the observed variables with the factors. We can obtain an estimate of $\mathbf{\Sigma}$ from data and then estimate some or all of the parameters on the right-hand side in order to get a better handle on how many and which variables are in the utility function for redistributive preferences.[2] However, none of the parameters on the right-hand side are identified (without further assumptions) unless $r < L(n) = 0.5\left(2n + 1 - \sqrt{8n+1}\right)$, where r is the number of latent factors, n is the number of observed variables, and $L(n)$ is known as the Ledermann (1937) bound.

Of course, factor analysis traditionally requires some assumptions that researchers are or should be uncomfortable making. But new methodological research in Goodrich (2008, 2009*a*,*b*) shows that some of

[2] In this paper, we utilize a polychoric correlation matrix, which is appropriate for ordinal variables. All of the variables discussed in the next section are ordinal except the individual's income. Thus, to fill out the correlation matrix, the polyserial correlation, which is appropriate when one variable is continuous and the other is ordinal, is calculated between log-income and the other variables.

these assumptions can be substantially relaxed in important ways. For example, factor analysis traditionally requires the researcher to specify in advance how many factors there are and to specify a distribution for the data, which is almost always the multivariate normal distribution. The estimator in Goodrich (2009*a*) requires neither of these assumptions, and a third difference follows from the first two: Rather than assuming that a model with a given number of factors is correct and demanding a statistically significant test statistic to reject this null hypothesis, we start with a null hypothesis that $r \geq L(n)$ (i.e. an identified factor analysis model does *not* fit) and seek a test statistic that allows us to reject this null hypothesis.

Goodrich (2009*b*) demonstrates this test has very good power, even with samples that are one-third the size of a typical nationally-representative survey. If this null hypothesis is rejected in favor of the alternative hypothesis that $r < L(n)$, then it is reasonable to conclude that a factor analysis model does hold in the population. Moreover, if $r < L(n)$, then the estimator in Goodrich (2009*a*) consistently estimates $\mathbf{\Theta}$, which allows additional quantities of interest to be estimated.

In summary, we need to analyze multiple political preferences in order to judge whether a particular theoretical approach is generally useful or can only explain a particular phenomenon. Factor analysis is a widely used tool in multivariate analysis, but the appendix follows Fessler (1996) to show that it is also appropriate for modeling utility-maximizing agents. The benefit of this approach is that we do not need to explicitly specify the functional form of the utility function or the inputs to the utility function. The primary cost is that the number of modeled outputs (n) needs to large enough relative to the number of common inputs (r); specifically r needs to be less than $L(n) = 0.5\left(2n + 1 - \sqrt{8n+1}\right)$.

4 Empirical Results

This section uses nationally representative, cross-country survey data to test the two competing theoretical approaches using the factor analytic estimator from Goodrich (2009*a*). First, we discuss the important features of the data before turning to the results.

4.1 Data

The data are taken from the 1999 International Social Survey Program (ISSP).[3] There are other waves of ISSP data and other organizations that collect cross-country data, but the 1999 ISSP data was chosen because it includes questions on both domestic and international redistribution. Germany is divided by the survey design into East and West, but we will refer to the following twenty-five places as "countries" anyway: Australia (AUS), West Germany (D-W), East Germany (D-E), United States (USA), Austria (A), Hungary (H), Norway (N), Sweden (S), Czech Republic (CZ), Slovenia (SLO), Poland (P), Bulgaria (BG), Russia (RUS), New Zealand (NZ), Canada (CDN), Philippines (RP), Israel (IL), Japan (J), Spain (E), Latvia (LV), France (F), Cyprus (CY), Portugal (P), Chile (RCH), and Slovakia (SVK). The sample sizes for each country are usually greater than 1000.

Most traditional political economy papers model the responses to the survey question below, where the responses are given on a standard five-point Likert scale:

> 7) How much do you agree or disagree with each statement about differences in income?
>
> B. It is responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes.

As was said before, traditional political economy regressions can only explain five to ten percent of the within-country variation in responses to questions like 7B. However, Ansolabehere, Rodden and Snyder (2008) argues that it important to model several survey items *simultaneously*, using techniques like factor analysis. In short, responses to each individual question can appear quite noisy, but are revealed to be less noisy when several survey questions are analyzed jointly.

One criticism of 7B is that it asks respondents how strongly they favor or oppose redistribution rather than how much redistribution they favor. In other words, it is possible that some respondent strongly believes that the government should give $1000 to everyone that earns less than $20,000, while another person favors a $5000 subsidy but is almost indifferent about it. The next question about progressive taxation provides an alternative that addresses this criticism somewhat.

[3] Analisis Sociologicos Economicos y Politicos, S.A.; International Social Survey Program, 2003-11-20, "International Social Survey Program (ISSP), 1985-2000", hdl:1902.29/CD-0158 Odum Institute;Zentralarchiv fur Empirische Sozialforschung and the Inter-university Consortium for Political and Social Research [Distributor]

8) Do you think people with high incomes should pay a larger share of their income in taxes than those with low incomes, the same share, or a smaller share? [much larger share, larger share, same share, smaller share, much smaller share]

In other words, the responses pertain to how progressive (or regressive) taxation should be rather than how strongly respondents feel about progressive taxation. Gans and Smart (1996) shows that the main individual-level result in the Meltzer-Richard model continues to hold under a progressive taxation scheme so these two questions are more-or-less equivalent from the perspective of traditional political economy.

In order to obtain leverage on the Fehr-Schmidt model, we need to some survey question that is *outside* the realm of traditional political economy. Kam, Cranmer and Fowler (2008) suggests that survey participants should play a DG as player 1 before the survey is administered, making it possible to operationalize social interest by the amount of money he or she chooses to give to player 2. This advice certainly would be expensive in large cross-country surveys, and in any event, cannot be done after-the-fact. However, we can approximately follow this advice with a critical survey question in the 1999 ISSP that pertains to *international* redistribution:

9) How much do you agree or disagree with each statement about international differences?

B. People in wealthy countries should make an additional tax contribution to help people in poor countries.

Although we would prefer that the ISSP ask how much the respondent believes people in wealthy countries should pay in additional taxes to help people in poor countries, 9B is still roughly analogous to a DG where Player 1 (the survey respondent) is a citizen of a reasonably wealthy country and Player 2 is a citizen of a poor country. Most of the twenty-five surveyed countries are wealthy (possibly excluding Chile and the Philippines) as far as 9B would be concerned.

Under a traditional political economy model, no one in a wealthy country would favor redistribution to a poor country and everyone in a poor country would favor it. That is, we argue that the existence of within-country variation in support for international redistribution is difficult to reconcile with the traditional political economy view. The key question becomes: How does the within-country variation in support for international redistribution covary with within-country variation in support for domestic redistribution and

progressive taxation?

To answer this key question while remaining open to the traditional political economy hypothesis, we include data on the respondent's self-reported social class, family income, and a "self-placement" question that asks "Where would you say you and your family actually are — at the top, somewhere in the middle, or toward the bottom [on a societal diagram]?". These variables are intended to capture where respondents fall in the productivity distribution and thus what is there private self-interest in a redistribution scheme. Unlike traditional political economy regressions where income is an explanatory variable that is implicitly assumed to be measured without error, here we model income, class, and self-placement as *outcome* variables that are stochastic functions of latent productivity and estimate the error variances.

Of course, the self-placement variable, the social class variable, and possibly even the questions on political preferences are extremely country-contextual, particularly when the data come from many advanced industrial countries, several Eastern European countries, and developing countries like Chile and the Philippines. Thus, we model each country separately and only use the *within-country* variation to draw conclusions. The samples for each country are so large that we do not need to do any pooling across countries in order to answer the fundamental questions raised at the beginning of this paper. In order to answer other, second-order questions — such as "Is the United States exceptional in its (lack of) inequality aversion?" — it would be necessary to pool across countries, perhaps in a multilevel model, and make additional, stronger assumptions to obtain confidence intervals for differences in estimates across countries. These second-order questions are nonetheless interesting and should be pursued in subsequent research once the question of whether theoretical models in the Meltzer-Richard family or the Fehr-Schmidt family are appropriate for a given set of countries.

4.2 Results

To reiterate, when $n = 6$ survey variables are modeled with factor analysis, the number of latent factors, r, to explain these outcomes could be any integer between zero and five. The traditional political economy view and the alternate view differ over the prediction that $r = 1$. The estimator derived in Goodrich (2009*a*) finds the minimum number of latent factors needed to explain the n observed variables, so if the traditional political economy view is correct that $r = 1$, then Goodrich's (2009*a*) algorithm would reach this

conclusion with certainty if given infinite data. Although our samples are not infinite, they are quite large, so our empirical conclusion that $r > 1$ and specifically that $r = 2$ in almost all of these countries is highly unlikely to be attributable to sampling variation.

If $r \leq 2$, then the error variances in the population are consistently estimated by the method in Goodrich (2009*a*). In a regression context, these estimates can be interpreted as $1 - \widehat{R}^2$. The traditional political economy literature has found that in regressions of preferences for domestic redistribution on income and other control variables that $1 - \widehat{R}^2 \geq 0.9$, which raises the uncomfortable dilemma that preferences are either fundamentally noisy or that the traditional political economy regressions are misspecified. The alternative camp takes the latter view and essentially predicts that the error variance in predicting individual preferences would decrease substantially if the empirical model controlled for the individual's inequality aversion. We estimate that the that the error variances in predicting preferences for redistribution are only half to two-thirds of the total variance, which goes against the traditional political economy view.

If the error variances are identified, then $\mathbf{\Sigma} - \mathbf{\Theta} = \mathbf{\Lambda}\mathbf{\Upsilon}\mathbf{\Lambda}'$ is identified. However, neither $\mathbf{\Upsilon}$, which is the correlation matrix among the r latent factors, nor $\mathbf{\Lambda}$, which is the matrix of coefficients in a regression of the n observed variables on the r factors, are separately identified without additional assumptions. Once we establish that $\mathbf{\Theta}$ is identified, we do make some additional assumptions in order to identify $\mathbf{\Upsilon}$and $\mathbf{\Lambda}$. The traditional political economy literature predicts that in a linear model where the dependent variable is either preferences for domestic redistribution or progressive taxation, the coefficient for the effect of productivity will be large and the coefficient for everything else will be (nearly) zero. The alternative camp disagrees and predicts that the coefficient for inequality aversion will differ substantially from zero, not only when preferences for domestic redistribution or progressive taxation is the dependent variable but also when preferences for *international* redistribution is the dependent variable. Our estimates support the alternative view.

Turning in more detail to the empirical results from the factor analytic estimator derived in Goodrich (2009*a*), in no country is the hypothesis that $r = 1$ supported by the data. Spain is the only country where we fail to reject the null hypothesis that $r \geq L(6) = 3$ on the basis of the test statistic, so $r = 2$ is a reasonable conclusion generally. Perhaps Catholic religiosity is an additional factor that is relevant in Spain but negligible in other countries. Unfortunately, neither Italy nor Ireland are included, but there is no

evidence of a third factor in Portugal or France. Since none of the parameters are identified in Spain without further assumptions, we do not consider Spain further in this paper.

Although we can reject the null hypothesis that $r \geq L(6) = 3$ in Latvia, Cyprus, Hungary, and the Czech Republic, there is evidence that raises a red flag as to whether a factor analysis model is appropriate for these four countries. Specifically, at least one of the estimated error variances is zero, which is both highly implausible for survey data and violates a technical assumption in made in the Appendix that no inequality constraints are active. Moreover, the results from these four countries are inconsistent with the results in the remaining twenty countries where a model with $r = 2$ factors yields an interior solution. Thus, we do not no consider these four countries further.

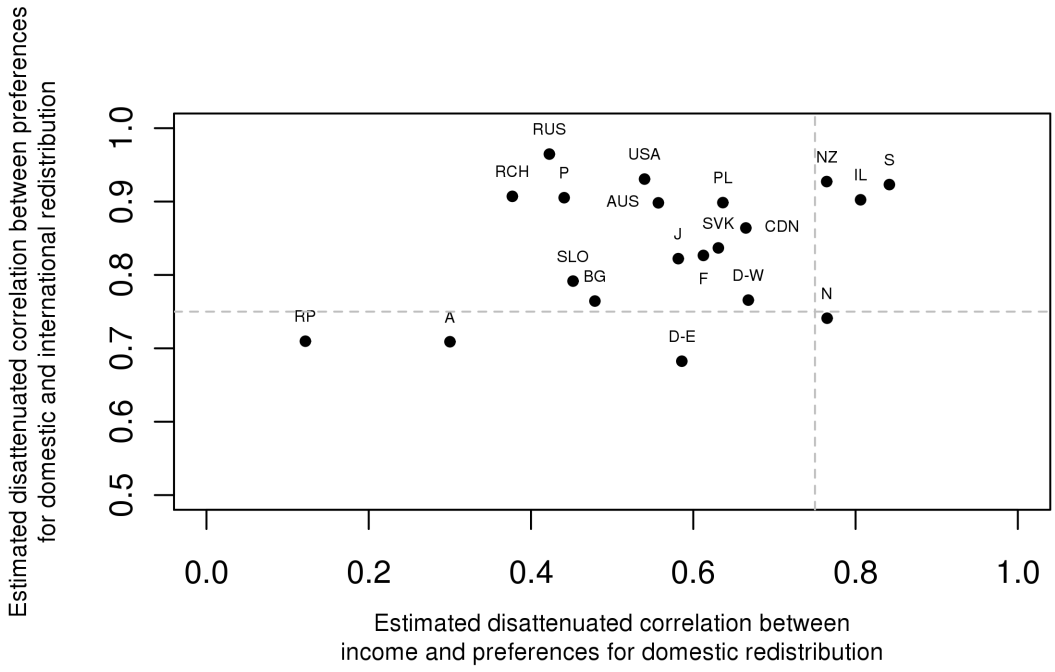
Many public opinion scholars would argue that many more than two explanatory variables — such as the respondent's sex, age, etc. — are needed to fully understand the variation in these six survey outcomes, but this criticism misses the point slightly. There may very well be a host of perhaps observable explanatory variables that affect these survey outcomes *through* these $r = 2$ factors, but two factors happen to be necessary and sufficient to render the survey responses conditionally uncorrelated. It is a worthwhile exercise to ask what these two factors are and what observable variables might drive them, but doing so requires stronger assumptions that are not always necessary to make the critical inferences in this paper.

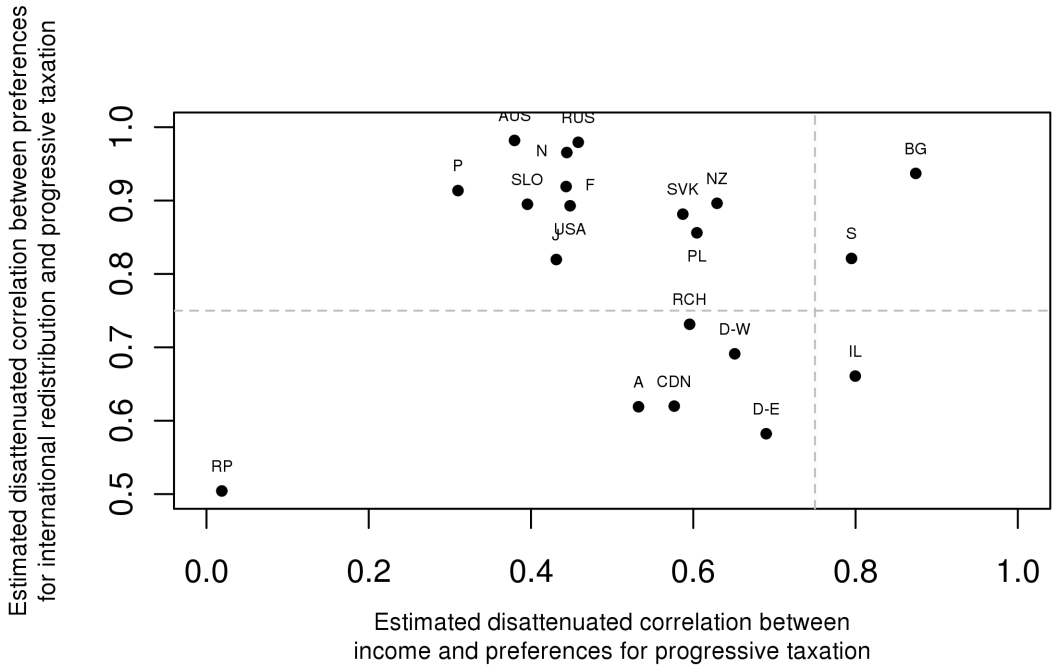
Given that the data are always inconsistent with the $r = 1$ hypothesis, the next question is whether the traditional political economy hypothesis can be sustained by showing that self-interest is sufficient for explaining preferences for domestic redistribution and progressive taxation. In short, this is not the case in these data. There are several pieces of evidence against the traditional political economy hypothesis.

First, we can analyze the estimated correlations between survey responses under the counterfactual that there were no random error (which is to say that we could perfectly measure income, preferences, etc.). Some disattenuated correlations are plotted in figure 1. If the traditional political model held, all the points would be concentrated in the southeast corner of the top half of figure 1, indicating a large, positive disattenuated correlation between preferences for redistribution and income and a near zero disattenuated correlation between preferences for domestic and international redistribution.

Instead we see in the top half of figure 1 that the average disattenuated correlation between preferences for domestic and international redistribution is about 0.84. In other words, excluding random error,

Figure 1: Estimated Correlations between Survey Questions under the "No Random Error" Counterfactual





Country abbreviations are listed on page 13.

preferences for domestic and international redistribution seem to have essentially the same data-generating process. Thus, if we take it as given that traditional political economy considerations are inadequate to explain within-country variation in preferences for international redistribution, they must also be insufficient to explain within-country variation in preferences for domestic redistribution. As a rough heuristic, we can speculate that self-interest drives a small wedge between preferences for domestic and international redistribution, rendering the disattenuated correlation between them somewhat less than 1.0.

This conclusion is reinforced by the average disattenuated correlation between income and preferences for domestic redistribution, which is only about 0.56. The pattern is qualitatively similar in the bottom half of figure 1, which plots disattenuated correlations involving preferences for progressive taxation. In the bottom half of figure 1, there is a downward-sloping relationship among the countries, which could be interpreted in multiple ways. However, the disattenuated correlations among policy preferences are still *much* stronger than can simply be attributed to the respondent's place in the income (really productivity) distribution.

Second, we can make two stronger assumptions that allow us to identify $\mathbf{\Lambda}$, the coefficients for the factors on each outcome variable. To identify the coefficients it is necessary and sufficient (see Reiersøl 1950) to assert two exclusion restrictions, one for each latent factor. The first is quite plausible, that stated income is not a function of inequality aversion, implying that stated income and self-interest differ only by a multiplicative constant and random error. The second assumption is more tenuous, but we assume that self-interest is unrelated to support for international redistribution. This assumption could be justified by a model where international redistribution is funded by a lump-sum tax on citizens in a wealthy country but would not hold exactly if taxation is proportional, progressive, or otherwise a function of wages.[4] Insofar as the question on international redistribution is a proxy for player 1's decision in a DG, it is suitable for identifying the effect of the second factor on the other outcomes, and we can interpret factor two like Fehr

[4]If everyone were exclusively motivated by private self-interest, then no one in a wealthy country would support international redistribution, but if taxes were an increasing function of income, then the rich in wealthy countries would be (even) more opposed to international redistribution than would be the poor in wealthy countries (assuming the marginal utility of a dollar is constant). Thus, the assumption that within-country variation is uncorrelated with individual productivity is valid under a traditional political economy model but only valid under the Fehr-Schmidt model if international redistribution were funded through lump-sum taxation. The experimental literature provides a further justification for this assumption. The contribution of player 1 in a DG is widely regarded as an important departure from private self-interest, although there is disagreement as to exactly how it should be interpreted. Kam, Cranmer and Fowler (2008) shows that player 1's contribution in a DG has predictive power for stated political preferences.

and Schmidt (1999) interprets β as an inequality aversion parameter in a DG. Unfortunately, there is little in the data that convincingly identifies α, although the fact that Goodrich's (2009) algorithm estimated that $r = 2$ suggests that α is not necessary in this battery of $n = 6$ survey questions.

Although it is possible that this exclusion restriction is not completely valid, if it is approximately true, then the bias in the other coefficients will be small enough to sustain inferences. If preferences for international redistribution are a function of the respondent's income to some extent, then we will likely slightly underestimate the effect of private self-interest on preferences and slightly overestimate the effect of inequality aversion. If these assumptions are badly wrong, then it is difficult to speculate about the size and direction of the bias.[5]

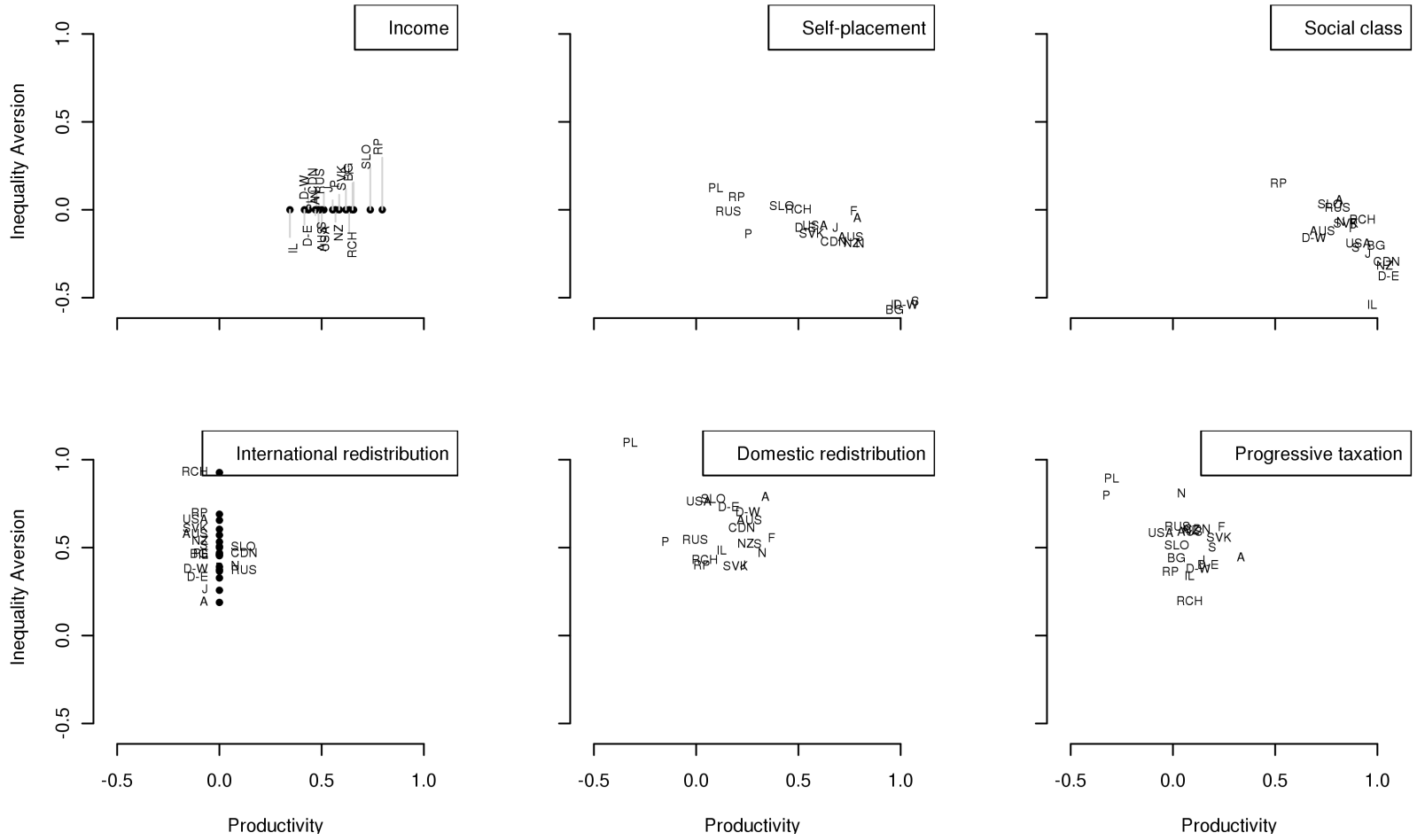
For the eighteen (mostly advanced industrial) countries where these assumptions seem reasonable, figure 2 plots the estimates of the free coefficients implied by these two exclusion restrictions. Stated preferences for domestic redistribution and progressive taxation are *not* exclusively a function of private self-interest, in contrast to the prediction of the traditional political economy models. This conclusion stems from the decidedly non-zero coefficients on the inequality aversion factor for these two survey questions.

Moreover, while stated preferences for domestic redistribution and progressive taxation in general are a function of both factors, the importance of inequality aversion is typically greater. One way to quantify the relative sizes of the effects is to counterfactually suppose that everyone within a country is forced to have the same value of a factor. If everyone had the same productivity, the average reduction in within-country variance in support for domestic redistribution would be about 8 percent. But if instead everyone had the same value of β, the average reduction in within-country variance in support for domestic redistribution would be about 37 percent. In short, the bulk of the variance in preferences for redistribution is attributable to variation in inequality aversion rather than variation in productivity. The additional explanatory power of the second factor is largely in harmony with the conclusion in Ansolabehere, Rodden and Snyder (2008) that apparent noise in individual survey questions is revealed to be consistency when multiple survey questions are modeled simultaneously.

It is tempting to try to make cross-country inferences from the results underlying figure 2, which is a

[5]In Poland and Bulgaria, there is evidence that one or both of these exclusion restrictions is invalid because they imply the coefficient of private self-interest on preferences for domestic redistribution has the wrong sign. Thus, we do not consider Poland and Bulgaria further.

Figure 2: Estimated Coefficients for Self Interest and Social Interest from the Factor Analysis Models



Each point represents a pair of estimated coefficients for a country, namely the coefficient for self-interest (productivity) and the coefficient for social-interest (β), for each of the six outcome variables. The coefficient for social-interest on family income is restricted to be zero, and the coefficient for self-interest on international redistribution is restricted to be zero. Each country is modeled separately, so these models do not permit statements about the statistical significance of cross-country differences. Country abbreviations are listed on page 13.

temptation that is mostly resisted here but would be interesting for future work. First, we are not modelling conditional means in support for domestic redistribution, progressive taxation, etc. Thus, we cannot say, for example, that the average Swede is more supportive of a policy than is the average American. We do model within-country covariance but do not pool across counties. In other words, the variation within each country is modeled separately, which precludes strict cross-country inferences about differences in estimates that come from separate models. But the point that should be taken away from figure 2 is that the results appear roughly similar across these eighteen countries.

One further point about the top-left panel of figure 2 is worth emphasizing. Stated income is measured with considerable error, as evidenced by the failure of the points to reach 1.0 on the x-axis. In the regressions that are commonplace in the political economy literature, the assumption is that all explanatory variables are measured perfectly. If this assumption fails to hold, the coefficients can be biased upward or downward, although in a very special case, namely OLS with exactly one covariate measured with random error, we know that the coefficient is attenuated. Thus, the traditional political economy regressions are afflicted with measurement-error bias, which is one reason that the proportion of explained variance is only five to ten percent. To circumvent this bias, researchers either need to instrument for income, which is difficult to do persuasively with survey data, or model the measurement error in the covariates explicitly, as is done in this paper. This methodological point pertains to survey research generally, not merely to the political economy of redistribution.

Furthermore, the estimated error variance in predicting social class is considerably *smaller* than the estimated error variance in predicting income (in proportional terms), which is to say that social class has more common variance with political preferences than does income. The distinctions between social class and income may seem minor, but this result runs counter to the traditional political economy finding that individuals whose utility is affected by private variables only set the marginal cost of redistribution equal to the marginal benefit. Social class is notoriously difficult to interpret, but we take it to be "subjective wealth", which is a "stock" variable rather than a "flow" variable like income. Traditional political economy models explicitly focus on income, rather than class or wealth, because income is the relevant variable at the margin. In the Meltzer-Richard model, someone who is wealthy but has below-average wages should *favor* redistribution (if fiscal policy is financed with taxes on wages). Conversely, class or wealth may be just as

important as income — if not more so — when considering the inequality that some people may be averse to. Overall, this finding that social class has more common variance with stated policy preferences is not conclusive but suggests that sociological interpretations of the bivariate association between prosperity and opposition to redistribution may be more applicable than the traditional political economy interpretations.

Some objections to the interpretation of these results are plausible but not ultimately fatal. Some would argue that the results favor a "rationalized self-interest" interpretation: the poor say they are inequality averse and the rich say they are inequality loving, but really both groups are just rationalizing their private self-interest. There could be some truth to this criticism, but an extreme version of this claim — that the second factor is unnecessary — is not supported by the data. If it were true that professed social interest were a simple, deterministic function of self-interest, then the method in Goodrich (2009*a*) would have found that only one factor is necessary and sufficient to explain the common variation in the survey outcomes. Nor is the extreme claim consistent with the estimated coefficients in figure 2. If the poor merely rationalize their private self-interest, why do they also tend to support international redistribution, which goes against the private self-interest?

It *is* the case that the two factors are somewhat correlated, which is possibly consistent with various theoretical stories. However, the estimated correlations between the two factors are moderate, with an average of about 0.41 across these eighteen countries. Hence, the *squared* correlations are not consistent with the extreme claim that all the variation in inequality aversion is merely rationalized private self-interest. An interesting avenue for future work is to investigate why the correlation between the two factors varies (considerably) across countries.

Others may prefer to interpret factor two as "propensity to state preferences that are inconsistent with private self-interest" due to the fact that survey responses are essentially costless for the respondent. Perhaps if a survey respondent were the decisive voter for a fiscal policy that greatly affected his or her disposable income, then the survey respondent would act in accordance with the traditional political economy hypothesis. Neither these survey data nor any data that social scientists could reasonably hope to gather completely eliminate this possibility.

The natural response is to test this hypothesis in a setting where individuals have to pay a tangible economic cost to redistribute, which has now been conducted hundreds of times in lab experiments, and the

overwhelming consensus is that private self-interest is insufficient to explain lab behavior, even when the economic stakes in the experiment are relatively high. This paper supports the external validity of these lab results to nationally representative samples. Also, it is consistent with Kam, Cranmer and Fowler (2008), where social-interest is operationalized by how California-representative respondents play in a DG, and this variable is found to predict stated political preferences in a post-survey.

A related criticism is that the wording of the question on international redistribution — which explicitly mentions "helping" poor countries — induces excessive stated support for international redistribution, given that actual international redistribution is tiny as a percentage of budgets. While many survey researchers have documented major question-wording effects, it is not clear that question-wording effects would have a major effect on the inferences drawn in this paper. To repeat, we are not modeling country-means; thus an additive shift in the distribution of responses due to the wording of the question would have no effect on the results. It is possible that the question-wording also reduces the variance in stated preferences for international redistribution, but we are not really modeling raw variances either, which are all normalized to be 1.0 because the latent variables assumed to underlie the ordinal response have no natural scale. Rather we are modeling within-country correlations. While it is possible that the wording of the question on international redistribution would have a major effect on the correlations with the other survey responses, it is not obvious that this would be the case, particularly because the question on international redistribution was asked after the questions on domestic redistribution and progressive taxation.

5 Conclusion

The distinguishing feature of traditional political economy models is that they assume a voter's utility depends only on private variables, such as the voter's (perhaps lifetime) income (distribution). Recent lab experiments have bolstered long-standing theories in sociology and psychology that individuals care about other people's payoffs. In particular, many subjects in experiments are willing to pay a personal cost in order to affect the distribution of payoffs. This paper has demonstrated that the same is true in nationally representative survey data for at least twenty countries. The data from some Eastern European countries were at odds with the model to varying extents, which is a puzzle that should be investigated more carefully, but the results for advanced industrial countries were quite consistent and largely unambiguous.

Many have argued empirically that traditional political economy models of preferences for redistribution are inadequate, but the results have had little effect on theorizing. The value-added of this paper is twofold. First, we use a dataset that allows us to simultaneously model stated preferences for domestic redistribution and international redistribution. Modeling a survey response that is outside the realm of traditional political economy — such as preferences for international redistribution — is critical to obtain leverage on whether traditional political economy theories of domestic redistribution preferences are adequate. Second, we use the recently-developed latent model variable from Goodrich (2009*a*), which makes weaker assumptions than other factor analysis estimators. In particular, it estimates the number of factors rather than assumes *a priori* that some number of factors are necessary and sufficient.

Taken together, these two innovations allow us to make stronger empirical claims. First, two explanatory variables are necessary and sufficient to explain the variation in these six survey outcomes in almost all countries (except Spain and possibly a few others). Conversely, we are able to reject the null hypothesis that three or more explanatory variables are necessary, although it is quite possible that observable socio-demographic characteristics of voters affect their preferences through these two factors, a possibility which should be studied further in subsequent work. Second, several bits of evidence point to the conclusion that preferences for domestic redistribution, progressive taxation, and international redistribution are produced by a qualitatively similar data-generating process. Since a traditional political economy model cannot easily explain within-country variation in preferences for international redistribution, we conclude that a traditional political economy model is also insufficient for domestic redistribution and progressive taxation.

The finding that private self-interest is insufficient to account for behavior in experimental distribution games has been the driving force behind inequality aversion theories for many years. The results in this paper suggest that these lab experiments have considerable external validity, so the political economy literature should expand its theoretical models, rather than persisting with the traditional assumption that utility only depends on private variables.

Moreover, researchers should attempt to control for inequality aversion when trying to demonstrate that any private variables explains some of the variation in preferences for domestic redistribution. Doing so persuasively is much easier in lab experiments than in most surveys, but the latent variable methods used in this paper illustrate how it can be accomplished if there is some variable — like support for international

redistribution — that is plausibly unrelated (or weakly related) to variation in productivity.

Unfortunately, these results imply that it will be difficult to show that preferences for redistribution can be ordered by a single, readily observable characteristic of the voters. Therefore, it will be difficult to prove that there is a Condorcet-winning level of redistribution that is dictated by the preferences of a recognizable decisive voter. There could easily be unholy coalitions of voters from across the productivity distribution that support or oppose changes in welfare-state policy due to the weight they attach to inequality aversion or differences in beliefs about how the policy change will affect inequality. This is a challenge and an opportunity rather than a fatal flaw in political economy.

The first task is to better understand how and why the two factors found in this paper are correlated and whether they may partially be a function of observable characteristics. The results in this paper suggest that the the degree of correlation varies widely across countries, but nothing in this paper suggests an explanation for this fact. The next step would be to integrate, extend, and test the aggregations mechanisms favored in the alternative literature, which rely more heavily on evolutionary game-theory and computational modeling than traditional political economy mechanisms like the single-crossing condition.

These are major undertakings, but the political economy literature on preferences for redistribution has not made that much *empirical* progress in the decades since Meltzer and Richard (1981). Of course, many traditional political economy papers have been written since then, but they have not greatly enhanced our ability to predict preferences for redistribution at the individual-level. The empirical results in this paper are consistent with those in Ansolabehere, Rodden and Snyder (2008) in the sense that policy preferences are not dominated by "noise"; rather there is systematic variance in preferences is not reducible to readily observable socio-demographic variables. Thus, if we want to make progress on modeling preferences for redistribution — or preferences for other policies that have major distributional impacts — we would be well advised to consider the systematic variance in preferences that is attributable to inequality aversion.

Appendix

This appendix shows that a factor analysis model is an appropriate tool when studying utility-maximizing agents. Let $\boldsymbol{\tau}_j^* = \begin{bmatrix} \tau_{j1}^* & \dots & \tau_{jk}^* & \dots & \tau_{jn}^* \end{bmatrix}'$ now be a n-vector of political preferences for voter j over

policy k and let $\boldsymbol{\eta}_j = \begin{bmatrix} \eta_{j1} \dots & \eta_{jp} & \dots & \eta_{jr} \end{bmatrix}'$ be a r-vector of characteristics of voter j whose union is relevant to j's political preferences. For example, in the Fehr-Schmidt model, $\boldsymbol{\eta}_j = \begin{bmatrix} Q(j) & \alpha_j & \beta_j \end{bmatrix}$. Under the utility-maximization framework, each voter chooses $\boldsymbol{\tau}$ to maximize some utility function, so $\boldsymbol{\tau}_j^* = \arg\max_{\boldsymbol{\tau}} U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)$.

The usual approach in the political economy literature is to assume that all voters have the same utility function so that differences in preferences are driven by differences in $\boldsymbol{\eta}$. A necessary condition for $\boldsymbol{\tau}_j^*$ to be optimal is the system of equations implied by the first-order conditions,

$$\begin{aligned} 0 &= \frac{\partial U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)}{\partial \tau_{j1}}, \\ &\vdots \\ 0 &= \frac{\partial U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)}{\partial \tau_{jk}}, \\ &\vdots \\ 0 &= \frac{\partial U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)}{\partial \tau_{jn}}, \end{aligned}$$

must simultaneously be satisfied.

In some cases, it is possible to obtain useful results without fully specifying the functional form of $U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)$, but rarely is it possible to obtain useful theoretical results without specifying what $\boldsymbol{\eta}$ is. One common choice is to assume that $\boldsymbol{\eta}$ is a scalar, η, capturing the single variable that is relevant for preferences. In the case of our traditional political economy model, $\eta_j = Q(j)$ is the only variable that drives variation in preferences for redistribution.

If there is a unique $\boldsymbol{\tau}_j^*$, then — even if we cannot write it explicitly — there is some vector function, $h\left(\boldsymbol{\eta}_j\right) = \begin{bmatrix} h_1\left(\boldsymbol{\eta}_j\right) & \dots & h_k\left(\boldsymbol{\eta}_j\right) & \dots & h_n\left(\boldsymbol{\eta}_j\right) \end{bmatrix}$, such that $\boldsymbol{\tau}_j^* = h\left(\boldsymbol{\eta}_j\right)$. If the functional form of $U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)$ is specified, then it is sometimes possible to derive $h\left(\boldsymbol{\eta}_j\right)$ explicitly by solving the first-order conditions to isolate each τ_{jk}^* on the left-hand side of the equation. However, closed-form solutions are typically only available when $U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)$ takes a simple form and $\boldsymbol{\eta}$ is a scalar or perhaps a vector with a very small length when there are additional constraints on the voter. For example, it is possible (but not shown) to obtain a closed-form solution for $\boldsymbol{\tau}_j^*$ in our traditional political economy model if productivity has

a Pareto distribution with carefully-chosen parameters. Thus, the convenience of the closed-form solution must be balanced against the strength and plausibility of the assumptions necessary to achieve it.

Fessler's (1996) alternative is to employ a first-order Taylor-series expansion, $\boldsymbol{\tau}_j^* = h\left(\boldsymbol{\eta}_j\right) \approx h\left(\overline{\boldsymbol{\eta}}\right) + \left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]\left(\boldsymbol{\eta}_j - \overline{\boldsymbol{\eta}}\right)$, where $\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]$ is a $n \times r$ matrix whose kth row and pth column contains $\frac{\partial h_k(\overline{\boldsymbol{\eta}})}{\partial \overline{\eta}_p}$. Thus,

$$\begin{aligned}
\operatorname{Cov}\left(\boldsymbol{\tau}_j\right) &= \mathbb{E}\left[\left(\boldsymbol{\tau}_j - \overline{\boldsymbol{\tau}}\right)'\left(\boldsymbol{\tau}_j - \overline{\boldsymbol{\tau}}\right)\right] = \operatorname{Cov}\left(h\left(\boldsymbol{\eta}_j\right)\right), \\
&\approx \mathbb{E}\left[\left(\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]\left(\boldsymbol{\eta}_j - \overline{\boldsymbol{\eta}}\right)\right)'\left(\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]\left(\boldsymbol{\eta}_j - \overline{\boldsymbol{\eta}}\right)\right)\right], \\
&= \left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]\mathbb{E}\left[\left(\boldsymbol{\eta}_j - \overline{\boldsymbol{\eta}}\right)'\left(\boldsymbol{\eta}_j - \overline{\boldsymbol{\eta}}\right)\right]\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]', \\
&= \left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]\operatorname{Cov}\left(\boldsymbol{\eta}\right)\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]'.
\end{aligned}$$

Fessler (1996) pushes this line of thought a bit further, and suggests a way to estimate $\left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]$ from knowledge of the objective function.

We can remain agnostic about the functional form of $U\left(\boldsymbol{\tau}; \boldsymbol{\eta}_j\right)$ by letting $\boldsymbol{\Sigma} = \operatorname{Cov}\left(\boldsymbol{\tau}\right)$, $\boldsymbol{\Lambda} = \left[\nabla h\left(\overline{\boldsymbol{\eta}}\right)\right]$, and $\boldsymbol{\Upsilon} = \operatorname{Cov}\left(\boldsymbol{\eta}\right)$. In that case, Fessler's (1996) approach implies $\boldsymbol{\Sigma} \approx \boldsymbol{\Lambda}\boldsymbol{\Upsilon}\boldsymbol{\Lambda}'$, which resembles a factor analysis model for $\boldsymbol{\Sigma}$ but assumes that each $\boldsymbol{\tau}_j$ is measured without error. In reality, at best we observe various proxies for $\boldsymbol{\tau}_j$ that contain unknown errors, which we hope are approximately uncorrelated with each other and with $\boldsymbol{\eta}$. If so, then $\boldsymbol{\Sigma} \approx \boldsymbol{\Lambda}\boldsymbol{\Upsilon}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is a diagonal matrix of measurement error variances, takes the classical form of a factor analysis model.

To briefly summarize Goodrich (2009*a*), given an estimate of $\boldsymbol{\Sigma}$ denoted $\mathbf{S}$, which is the estimated (polychoric) correlation matrix among responses to n survey questions, the method chooses $\widetilde{\boldsymbol{\Theta}}$, a proposal for $\boldsymbol{\Theta}$, to maximize the eigenvalue "dispersion" of $\mathbf{S} - \widetilde{\boldsymbol{\Theta}}$, subject to the restriction that $\mathbf{S} - \widetilde{\boldsymbol{\Theta}}$ is positive semi-definite (PSD). If instead $\boldsymbol{\Sigma}$ were available, then maximizing the eigenvalue dispersion would also minimize the rank of $\boldsymbol{\Sigma} - \widetilde{\boldsymbol{\Theta}}$, where the minimum rank, r, is equal to the minimum number of factors that are consistent with a factor analysis model for these n variables (see Thurstone 1935; Reiersøl 1950). Although $\boldsymbol{\Sigma}$ is never available, applying the method to $\mathbf{S}$ produces asymptotically equivalent results as the sample size becomes infinite, and if $r < L\left(n\right)$, then the method in Goodrich (2009*a*) consistently estimates $\boldsymbol{\Theta}$. With survey data, the samples are typically very large so sampling error is negligible.

References

Alesina, A.F. and P. Giuliano. 2009. "Preferences for Redistribution." *NBER Working Paper w14825* .

Ansolabehere, S., J. Rodden and J.M. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(02):215–232.

Corneo, G. and H.P. Grüner. 2000. "Social limits to redistribution." *American Economic Review* pp. 1491–1507.

Cusack, T., T. Iversen and P. Rehm. 2006. "Risks at Work: The Demand and Supply Sides of Government Redistribution." *Oxford Review of Economic Policy* 22(3):365.

Fehr, E. and K.M. Schmidt. 1999. "A theory of fairness, competition, and cooperation*." *Quarterly journal of Economics* 114(3):817–868.

Fessler, J.A. 1996. "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography." *IEEE Transactions on Image Processing* 5(3):493–506.

Fong, C. 2001. "Social Preferences, Self-Interest, and the Demand for Redistribution." *Journal of Public Economics* 82(2):225–46.

Gans, J.S. and M. Smart. 1996. "Majority voting with single-crossing preferences." *Journal of Public Economics* 59(2):219–237.

Georgiadis, A. and A. Manning. 2007. "Spend it like Beckham? Inequality and redistribution in the UK, 1983-2004.".

Gilens, M. 2000. *Why Americans hate welfare: Race, media, and the politics of antipoverty policy*. University of Chicago Press.

Goodrich, B. 2008. "Semi-Exploratory Factor Analysis and Software to Estimate It." Paper presented at the annual meeting of the APSA 2008 Annual Meeting, Boston, MA.

Goodrich, B. 2009*a*. "Bringing Rank-Minimization Back In: Estimating the Number of Common Inputs to a Data-Generating Process." Unpublished paper, Harvard University.

Goodrich, B. 2009*b*. "Choosing the Number of Latents with the Indirect Rank-Minimization Algorithm." Unpublished paper, Harvard University.

Kam, C.D., S.J. Cranmer and J.H. Fowler. 2008. "When It's Not All About Me: Altruism, Participation, and Political Context." Unpublished paper, available from http://jhfowler.ucsd.edu/.

Ledermann, W. 1937. "On the rank of the reduced correlational matrix in multiple-factor analysis." *Psychometrika* 2(2):85–93.

Lübker, M. 2007. "Inequality and the demand for redistribution: are the assumptions of the new growth theory valid?" *Socio-Economic Review* 5(1):117.

Meltzer, A.H. and S.F. Richard. 1981. "A Rational Theory of the Size of Government." *The Journal of Political Economy* 89(5):914–927.

Reiersøl, O. 1950. "On the identifiability of parameters in Thurstone's multiple factor analysis." *Psychometrika* 15(2):121–149.

Thurstone, L.L. 1935. *The vectors of mind: multiple-factor analysis for the isolation of primary traits*. The University of Chicago Press.