Majoritarian Electoral Systems and Consumer Power: A Matching Rejoinder*

Daniel E. Ho†

September 1, 2003

Abstract

In a substantial contribution to positive political economy, Rogowski and Kayser (2002, p. 526) finds that "systems of proportional representation . . . systematically advantage producers and disadvantage consumers." I find that there is no evidence to sustain that conclusion. The original study extrapolates severely due to the fact that proportional representation countries are systematically different from majoritarian district countries in their background characteristics. Accounting for these differences by matching on the propensity score forces us to discard observations that severely extrapolate from the data, yielding estimates of such high variance that we cannot find evidence for the price-level effect. The only way to reassess the hypothesis without bias or implausible assumptions is to gather a larger dataset, which by including non-OECD democracies increases potential observations. Yet even with this data the price-level effect remains undetectable. The conclusion of a price-level effect thereby rests on modeling assumptions that are theoretically and empirically unjustified.

*I thank Ronald Rogowski and Mark Kayser for helpful discussions on earlier versions of this paper and for graciously providing their data. Gary King provided many constructive critiques and suggestions. Thanks also to Jim Alt, Ian Ayres, Alexis Diamond, José Fernández-Albertos, Jeffry Frieden, Kosuke Imai, Dale Jorgenson, Sophie Langenskiöld, Kevin Quinn, Ken Shepsle, Jiyeoun Song, Elizabeth Stuart, and participants of the 2003 Annual Meeting of Political Methodology and the Research Workshop in Political Economy at Harvard University for helpful discussions and comments. Research support was provided by the Center for Basic Research in the Social Sciences. All errors are mine alone. Software and data generated for this study will be available at http://www.people.fas.harvard.edu/˜deho.

†Ph.D. Candidate, Department of Government, Harvard University; J.D. Candidate, Yale Law School; Center for Basic Research in the Social Sciences, 34 Kirkland Street, Cambridge MA 02138, Tel: 617-384-9313, Fax: 617-496-2254, http://www.people.fas.harvard.edu/˜deho, deho@fas.harvard.edu.

1 Introduction

In a substantial contribution to positive political economy, Rogowski and Kayser (2002, p. 526) (RK) finds that "systems of proportional representation (PR) systematically advantage producers and disadvantage consumers." If true, this previously unnoticed welfare effect of electoral systems might not only help to explain such puzzles as the "exceptional...frequency of electoral reform" in the 1990s (Perrson and Tabellini 2003, p. 77) but would also seem to defy the complex systematization of electoral rules by political scientists such as Cox (1990) by its reduction of election rules into two discrete and ideal types (majoritarian and PR). Clearly, the validity of the study bears significant policy and intellectual consequences, as evidenced by the wide reception of the RK study in the field.[1] This paper concerns itself with the *empirical* validity of the RK claim.

1.1 The Causal Logic

The intuition of the causal logic in the formal model developed by RK is as follows: First, employing a Stigler-Peltzman model of regulation, politicians maximize political support which is a function of money and votes. While producers can provide both money and votes, consumers can only vote to influence political actors. Politicians thereby trade off support from producers and consumers, equating the marginal rate of substitution between the two groups.

Second, the seats-votes elasticity, defined as "the percentage increase in seats to be anticipated from a one percent increase in votes" (p. 530), is assumed to be greater in majoritarian systems than in PR systems. That is, since the marginal vote in a majoritarian system can more drastically change the allocation of seats, politicians should cater more towards consumers in majoritarian as opposed to PR systems. Intuitively, the marginal difference between 49% and 51% of the vote share in a majoritarian system can *determine* the winner, whereas in a pure PR system, the additional 2% of the vote simply allocates more legislative seats *proportionally*.[2] Therefore, in a majoritarian system politicians will allegedly care more about the marginal

[1] Evidence of the impact of the article abounds. In the short time since the original article was published, Scartascini (2002) has directly extended the Rogowski-Kayser model to examine business regulation; Milner and Judkins (2002) has found supporting evidence in the effect of majoritarian electoral systems on policies that affect price levels, specifically examining protectionism; Rosenbluth and Schaap (2002) has found empirical evidence on the effect of electoral rules on banking regulation; and Fiona (2002) analyzes inter-industry stock market price variation. The article has been widely cited, e.g., Gourevitch, Carney and Hawes (2003), Gourevitch (2003), Scartascini and Crain (2002), Pablo T. Spiller and Tommasi (2003), and Iversen and Soskice (2002) describe the tension between simple median voter frameworks and, *inter alia*, the Rogowski-Kayser findings as the "electoral system puzzle." Iversen and Soskice (2002, p. 2). Moreover, Rogowski, Chang and Kayser (2002) expands the empirical tests of the price-level effect to a time-series cross-sectional study spanning OECD countries from 1970-2000. However, the problems of extrapolation and causal inference discussed in this paper are arguably even more severe in a panel setting.

[2] Of course in a PR system the number of seats in a district determines whether just *how* proportional the vote gain is translated into seats.

	Specification		
	Model 1.1	Model 1.3	Model 1.4
Effect of SMD on Prices	-13.99^{**}	-12.40^{**}	-10.45^{*}
	(4.71)	(3.83)	(4.96)
Controls Included			
GDP per capita	*Yes*	*Yes*	*Yes*
Trade Openness	*No*	*Yes*	*Yes*
3 Year Exchange Rate Appreciation	*No*	*Yes*	*Yes*
Log of Arable Land / Population	*No*	*No*	*Yes*
Log of Population	*No*	*No*	*Yes*
Log of Energy Production / Consumption	*No*	*No*	*Yes*
N	24	22	22

Table 1: Summary Rogowski-Kayser findings in OLS regression analysis from Rogowski and Kayser (2002, p. 533). Standard errors in parentheses. *$p < 0.1$; **$p < 0.01$. Yes/No indicates whether variable was included in regression model.

voter, or the consumer.[3]

Third, RK posits that price levels, or more accurately deviations from competitive world prices, reflect the tradeoff between producers and consumers, with consumers favoring lower prices. Therefore, majoritarian electoral systems systems should systematically favor consumers by exhibiting lower price levels than PR countries.

1.2 Assessing the Price-Level Effect

Accordingly, RK uncovers a never before observed effect in their empirical study: single-member district (SMD) electoral systems lead to a decrease in roughly 12% of national prices, plus or minus 7%, compared to proportional representation systems (hereinafter, the "price-level effect") (Rogowski and Kayser 2002, p. 533, Model 1.3, using 95% confidence interval). Table 1 summarizes their main results and model specifications using linear regression. In their own words, "[t]he clear finding is that – controlling for virtually every other relevant influence – prices of goods and services are systematically higher in PR countries" (p. 526).

What RK overlooks is the model-dependency of the price-level effect due to the fact that countries with single-member district electoral systems are *systematically different* from countries with proportional

[3]Note that this claim relies on an assumption of a two-party competition with both parties receiving close to 50% of the vote. It is certainly possible that in a two-party competition where one party *ex ante* expects a much larger vote share than the competitor, the seats-votes elasticity would actually be lower for a majoritarian than a PR system. This is easily verified by examining the graph of the seats-votes elasticity in note 15, p. 531, of Rogowski and Kayser (2002) (replicated in Figure 7 of the Appendix). In their stylized example, PR systems are more pro-consumer when the vote of any one candidate exceeds 0.684 (i.e., to the right of the intersection between the PR and SMD curves).

representation systems in background characteristics that affect price levels.[4] SMD countries import substantially less goods as a proportion of GDP, are geographically much larger, have far more arable land, and have significantly higher ratios of energy production to consumption. Despite the fact that these variables are included in regression analysis, the imbalance in these characteristics between SMD and PR countries makes any inference about the causal effect of electoral systems highly model-dependent.[5] Any inference about the price-level effect may face the problem of extrapolating too far from the data. Intuitively, it may be very difficult to assess what the price levels of PR countries would be *if* they were SMD countries, given that SMD and PR countries are so drastically different.

RK is aware of the potential sensitivity of the prive-level effect, and conducts examplary robustness tests using casewise diagnostics (p. 536-37). Yet a more direct technique to assess the robustness of causal effects given systematic differences between SMD and PR countries, namely propensity score matching, might have allowed RK to perform a more exhaustive analysis of the price-level effect. Propensity score matching has been widely applied to problems of causal inference in various disciplines with non-randomized observational studies (see, e.g., Holland 1986, Rosenbaum and Rubin 1983b, Dehejia and Wahba 1999, King and Zeng 2002, and Imai 2003). While it was not widely used in political science at the time, had the technique had been available to RK it would likely have led to a different conclusion as shown herein.

Section 2 employs a more formal definition of extrapolation to assess the severity of the problem in the RK data. Section 3 provides an overview of the problem of causal inference as embodied in the comparative study of electoral institutions, and provides a brief formal rationale for using propensity score matching to assess the effect of SMD systems. Applying this framework, Section 4 finds that RK's claim of a "significant and robust" price-level effect of majoritarian electoral systems is incorrect (Rogowski and Kayser 2002, p. 526). Matching exposes a bias-variance tradeoff arising from the small number of comparable SMD and PR countries in the RK data. Section 5 therefore expands the original dataset from OECD countries to all democracies, but again finds little support for the broad claim of price-level effect. Section 6 concludes.

2 Extrapolation in the RK Data

Extrapolation is often informally described as an inference made outside of the available data range (Manski 1995, pp. 1-20). For example, suppose we are interested in the effect of trade openness on price levels. First consider an inference made *inside* of the available data range, or an interpolation (see King and Zeng 2002).

[4] PR countries in their dataset are Spain, Norway, Belgium, Turkey, Germany, Japan, Sweden, Greece, Denmark, Ireland, Portugal, Netherlands, Austria, Finland, Italy, and Switzerland. SMD countries are Australia, Canada, France, New Zealand, the United Kingdom, and the United States.

[5] By imbalance, I simply mean that PR and SMD countries are not comparable along these dimensions.

RK's Model 1.2 would suggest that an increase of 1% in Portugal's trade openness would lead to a decrease of 0.5% in Portugal's price levels, plus or minus 0.4%. This qualifies as an interpolation since there are countries with similar levels of trade openness to approximate the counterfactual. An inference based on extrapolation, however, would suggest that an increase in 100% in Portugal's trade openness is associated with a decrease of 50% in Portugal's price levels, plus or minus 80%. While the regression model would easily provide this first difference calculation, no countries with such high values of trade openness exist in our data to allow us to examine this counterfactual. Any inference based on such similar counterfactuals extrapolates too far, relying critically on model-dependent assumptions. There may, for example, exist declining marginal price effects for extremely open economies, which would indicate that we overestimated the first difference by assuming a linear additive effect. While model-based variance estimates do increase slightly as we extrapolate, this is *conditional* on the functional form assumptions of the model that identify a quantity of interest. Relaxing these assumptions will increase the variance of the quantity of interest, forcing us to examine only the *informative* data that identifies a quantity of interest without extrapolation (i.e., interpolation).

To define extrapolation more formally, let X denote an observed data matrix of explanatory covariates, and x a vector of values of the covariates, in this case representing characteristics of a particular country. Anytime that x is not a row contained in X, we are making a counterfactual inference, common to most if not all comparative statics claims. King and Zeng (2002, p. 7) define extrapolation stating that "questions that involve interpolation are values of the vector x which fall in the *convex hull* of X," where the convex hull is the smallest convex set containing a set of points.

To provide some intuition, suppose we simply had two variables, GDP per capita and price levels in the RK data. Figure 1 depicts the convex hull of this set of points, represented by the gray-shaded polygon with vertices at extreme points of the data, where each observed data point is labelled by the country abbreviation. The graph also provides simple fitted curves for a linear and quadratic specification, regressing prices on GDP per capita. The black squares, denoted by *A*, *B*, and *C*, represent counterfactuals. Suppose we were interested in the price level of a country with a GDP per capita of US $15,000, represented by *A*. Using both linear and quadratic models we would roughly estimate that this country would have price levels of 101, roughly equal to that of the United States. Since *A* falls within the convex hull, it qualifies as an interpolation. Now consider our estimation for counterfactual of a country with GDP per capita of US $40,000, represented by *B* and *C*. This counterfactual clearly falls outside of the convex hull and is thereby an extrapolation. More importantly, we can see that the model specifications yield price estimates that differ

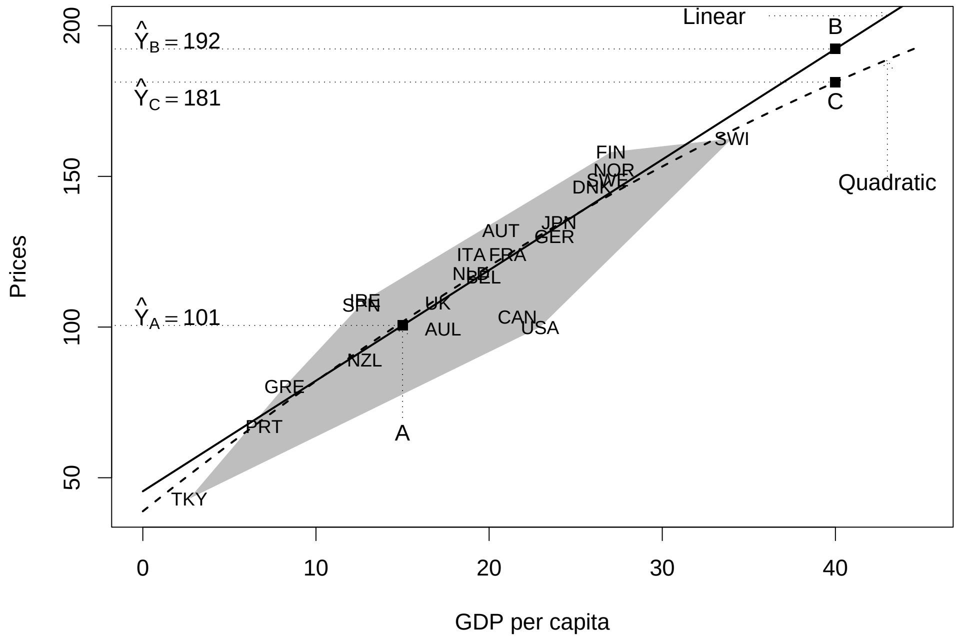

Figure 1: The gray polygon represents the convex hull of the data consisting of GDP per capita and price levels in the RK dataset. All countries are labelled. Dashed lines represent OLS estimations using a linear ($\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 GDP$) and quadratic ($\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 GDP + \hat{\beta}_2 GDP^2$) specification. Squares A, B, and C represent counterfactuals and $\hat{Y}_A$, $\hat{Y}_B$, and $\hat{Y}_C$ their predicted values.

by roughly 11%, the vertical distance between B and C, despite the fact that the in-sample fit of these models is virtually identical. This illustrates that extrapolations rely to an extent far greater than interpolations on model specification. Linear and quadratic predictions diverge the more we extrapolate.

Similarly, the counterfactual of what the price levels of countries would be if PR countries were SMD countries relies on the existence of *comparable* SMD countries. Without comparable countries, the counterfactual extrapolation relies too critically on modeling assumptions. Note that there is no easy way to get around counterfactual reasoning: in order to learn about the effect of SMD systems on price levels, we use countries that are the same in all respects except for the electoral system.

To illustrate extrapolation in a simple setting, Figure 2 provides the density estimates (smooth versions of histograms) of SMD and PR countries for GDP per capita on the left panel x-axis and arable land on the right panel x-axis. The solid curves represent SMD countries and the dashed curves represent PR countries. In the left panel, note that PR countries have a far wider range of incomes. Unlike PR countries, there are no SMD countries with GDP per capita above US $30,000 or below US $8,000. An inference about the counterfactual of changing an SMD country to PR outside of this range where the densities overlap is an

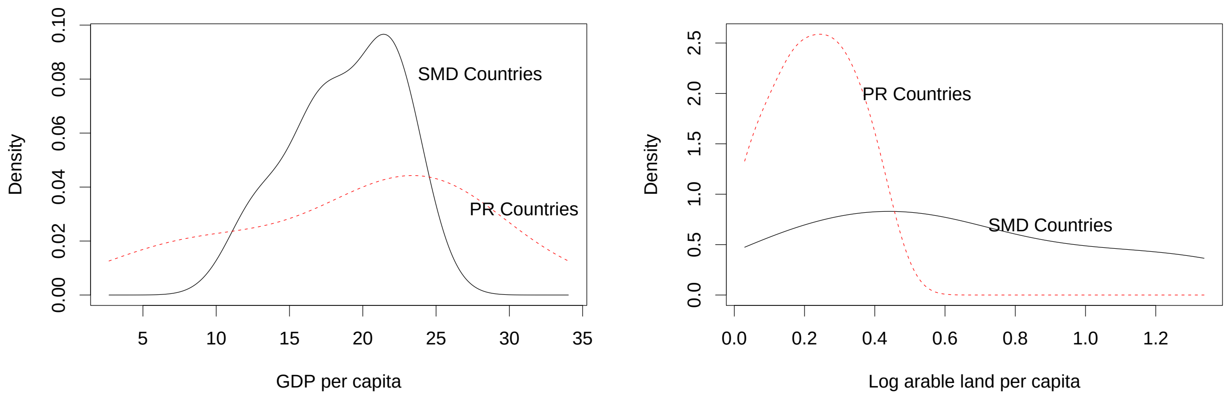

Figure 2: Density estimates of GDP per capita and arable land SMD countries (solid curve) and PR countries (dashed curve). In the left panel, the counterfactual of changing SMD to PR countries is outside of the convex hull for GDP per capita above US$ 30,000 and below US$ 8,000. In the right panel, the counterfactual of changing PR to SMD countries is outside of the convex hull above arable land of roughly 0.4.

extrapolation. The right panel demonstrates that SMD countries have significantly more arable land per capita than PR countries. A counterfactual beyond arable land of roughly 0.4 constitutes extrapolation.

Consider now the implications for extrapolation in the latter instance of arable land. Figure 3 presents fitted values for two linear regression models using only arable land and SMD as variables. The left graph presents a reduced RK specification, where the electoral system is modelled as a dummy variable. The solid lines represent fitted values for SMD countries and the dashed lines represent fitted values for PR countries. With this specification we would estimate roughly an 8% price decrease at all values of arable land – simply the difference in the two parallel lines. While this model assumes linearity and additive treatment effects across the full range of arable land, note that there are only two SMD countries, signified by a box around the country label, that are within the range of arable land for PR countries, signified by the country labels without boxes: only France and the United Kingdom are SMD countries that have arable land levels below 0.4. Any inference from our data about price levels of PR countries with *high* levels of arable land is highly model-dependent. The right panel presents the fitted values using separate regressions for PR and SMD countries. Note how the fitted lines now intersect roughly at arable land values of 0.45. In this specification, we would predict that at high ranges of arable land SMD countries have higher prices, wheras at low ranges of arable land SMD countries have lower prices than PR countries. In other words, the change in specification would lead to a complete reversal of the causal effect at high levels of arable land! Clearly, we should be cautious about extrapolation here. Note again here that interpolations are not as sensitive to the specification change – the estimate of the treatment effect within the range of overlapping values of arable

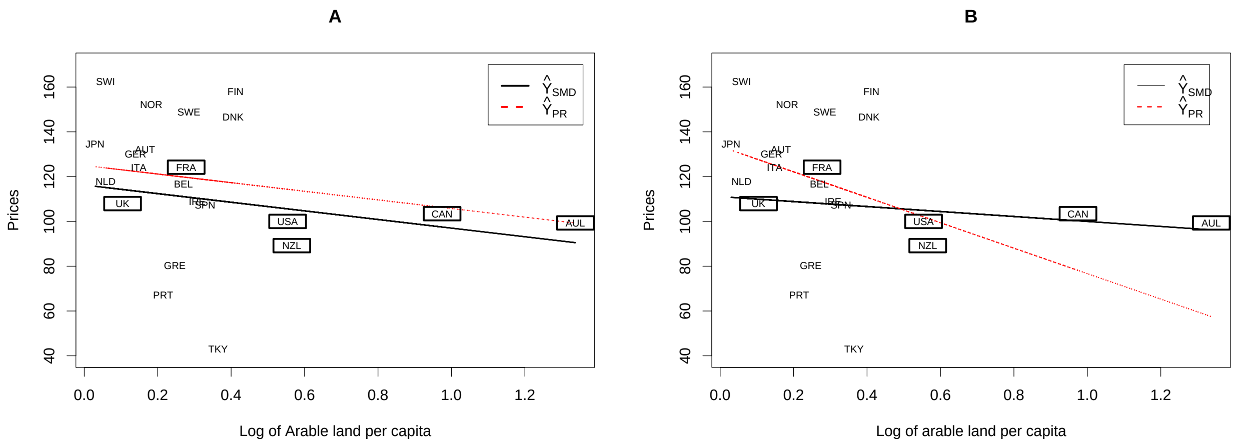

Figure 3: OLS fitted values for (a) model regressing prices on arable land and dummy variable of SMD (left panel) and (b) model regressing prices on arable land separately for SMD and PR countries (right panel). Solid lines represent fitted values for SMD countries and dashed lines represent fitted values for PR countries. Squares around country labels represent observed SMD country.

land is quite similar.

While such simple two-dimensional examples illustrate the intuition of extrapolation analysis, a geometric visualization of the convex hull and the problem of extrapolation becomes more difficult as we increase the number of explanatory covariates. I employ King and Zeng's methods to calculate how many counterfactuals for the full set of covariates in RK's study are in the convex hull.[6] Each counterfactual is created by simply changing the electoral system from its observed value to its counterfactual (e.g., SMD countries to PR, or PR countries to SMD), while holding all other covariates at their observed values. Table 2 presents the results of this analysis. Even in RK's Model 1.1, which simply includes GDP per capita as a control, the majority of the counterfactuals are extrapolations. As Figure 2 illustrated, all six SMD countries are in the region of overlapping densities for GDP per capita whereas extreme values of the PR countries are not. Table 2 clearly reflects this, as 100% of the counterfactuals changing SMD countries to PR countries are in the convex hull, whereas only 28% of the counterfactuals changing PR countries to SMD countries are. This suggests that the data might better be able to explain the effect of SMD on SMD countries than on PR countries. If we define SMD as a treatment, this would be referred to as the treatment effect on the treated. Extrapolation becomes quickly more severe with more fully specified models. In Model 1.3, only 18% of the counterfactuals are inside the convex hull. Even worse, in Model 1.4 *none* of the counterfactuals are inside the convex hull of the data. In other words, in the fully-specified model, every counterfactual inference is an extrapolation! The consequences of such extrapolation are demonstrated by simple changes in the model

[6]The software by King and Zeng to solve this linear programming problem is available at `http://GKing.Harvard.Edu`.

		Number of counterfactuals in convex hull					
		SMD → PR		PR → SMD		Total	
Model	N	No.	%	No.	%	No.	%
RK 1.1	24	6	100	5	28	11	46
RK 1.3	22	3	50	1	6	4	18
RK 1.4	22	0	0	0	0	0	0

Table 2: Number and percentage of counterfactuals falling within convex hull of data. SMD (PR) → PR (SMD) indicates a counterfactual of changing an SMD (PR) country to PR (SMD) while holding all other covariates at their observed values. RK refers to Rogowski-Kayser specifications.

specification, the traditional way to conduct sensitivity analyses of model-dependent conclusions. For example, in Model 1.4, the substitution of a squared term for GDP per capita yields an insignificant coefficient for SMD (p=0.13). Similarly, adding a squared term for energy leads to an insignificant coefficient for SMD (p=0.13). Or the addition of an interaction term between SMD and energy results in no substantive effect of the electoral system. (See Appendix, Figure 8.)

Yet of course not all extrapolations are equally bad. An extrapolation just outside of the convex hull that has many data points close by may even represent a better inference than an interpolation that has an isolated datapoint bounding the convex hull. To go back to our earlier example of income and price levels, Switzerland has the highest income in the RK dataset, with a GDP per capita of roughly US $34,000. A counterfactual of what its price level would be if its income increased to US $35,000 is obviously a better inference than if its income increased to US $100,000, yet both are outside the convex hull. Therefore, we turn to a framework of causal inference that explicitly takes into account degrees of differences between SMD and PR countries.

3 Assessing the Causal Effect of Electoral Systems

3.1 The Causal Effect

A cause is a intervention that leads to some observable effect in our outcome variable as compared to no treatment (i.e., control). More formally, let i index the countries of our interest. Let $Y_i(1)$ represent the potential outcome of interest when the country is SMD and $Y_i(0)$ represent the potential outcome when the country is PR. The treatment effect for country i is defined as $Y_i(1)-Y_i(0)$. The average treatment effect for the treated population is defined as:

$$E(Y_i(1)|T_i=1)-E(Y_i(0)|T_i=1),$$

Variable	SMD Mean	PR Mean	T-Statistic
Per Capita GDP	18.81	19.41	0.23
Trade Openness	0.07	-0.01	-1.67
3 Year Exchange Rate Appreciation	6.68	-6.48	-0.98
Log of Arable Land / Population	0.64	0.23	$-\mathbf{2.18}$
Log of Population	10.4	9.55	-1.29
Log of Energy Production/Consumption	0.88	0.51	$-\mathbf{2.44}$
Imports as a proportion of GDP	0.18	0.30	**2.93**
Size (1,000 km^2)	4,754	244	**-2.26**

Table 3: T-statistics comparing single-member district countries (SMD) and proportional representation countries (PR) on selected covariates. Bold signifies $|p| < 0.10$. $N = 22$.

where T_i signifies whether country i was assigned treatment. The notation in this section largely trails that of Dehejia and Wahba (2002) and is consistent with the standard literature on causal inference (see, e.g., Holland 1986 and King, Keohane and Verba 1994). The fact that we cannot observe a control country when it has been assigned treatment, $Y_i(0)|T_i = 1$ is known as the fundamental problem of causal inference (Holland 1986, p. 947, King, Keohane and Verba 1994, p. 79). Concretely, we can never observe the counterfactual of a country under a PR system when it in fact has been "assigned" a majoritarian system, or the counterfactual of a country under a majoritarian system when it in fact has been "assigned" a PR system. In experimental studies, randomization provides a way of estimating the treatment effect by balancing all countries on characteristics and thereby:

$$E(Y_i(0)|T_i = 1) = E(Y_i(0)|T_i = 0).$$

Randomization therefore plays the central role of balancing all other covariates. By randomly assigning countries a majoritarian electoral system, the effect of other factors (e.g., geography, GDP per capita, imports/GDP) balances out, and the difference that remains can be attributed to the treatment.

3.2 Majoritarian Versus PR Countries

Table 3 depicts simple means comparisons between PR and SMD countries. SMD countries are strikingly different from PR countries. Import penetration is 60% lower on average in SMD countries than in PR countries (p=0.008), and SMD countries on average trade 8% less as a proportion of GDP, though this is only borderline significant. Moreover, it appears that PR countries have significantly less arable land per population and have lower energy productions, both at significant levels (p=0.078 and 0.027, respectively). Such differences make it extremely difficult to assess what the causal effect of the electoral systems might be – as shown earlier, inferences based on OLS extrapolation are highly model-dependent.

4 Matching on the Propensity Score

Given that differences between SMD and PR countries exist on multiple dimensions, our goal is to match countries on observable characteristics so as to obtain an unbiased estimate of the treatment effect. The intuition is that we only care about countries that are closely aligned in observable characteristics such as import penetration and factor endowments. The Netherlands, Greece, and Switzerland, for example, have far higher import penetration levels than all SMD countries. Using such countries that are highly unlikely to have SMD systems to estimate $Y_i(0)|T_i = 1$ may lead to significant bias. In short, the key objective is to assess the causal effect by examining countries comparable in all covariates but the electoral system.

In the simple case of one additional explanatory variable x_i, the intuition is simple: we match PR and SMD countries with similar values of x_i and the difference in outcome variables is our estimate of the treatment effect. Matching on the host of covariates in RK's models, however, proves difficult when exact matches are unavailable since it is unclear which covariates matter and how to weight them. Each additional variable leads us down to the curse of dimensionality, namely that each additional variable increases the number of parameters required to estimate $E(Y|X)$ geometrically in the number of covariates.

Therefore, to summarize these differences, I calculate the propensity score, $e_i(X_i) = Pr(T_i = 1|X_i)$, which represents the probability that a country has an SMD system conditional on other covariates X_i. The propensity score is akin to a one-dimensional summary of all relevant covariates, and Rosenbaum and Rubin (1983b) shows in a seminal result that conditioning on the propensity score is bias reducing in the same degree as conditioning on all the covariates. Matching on the propensity score thereby enables us to estimate the treatment effect, where:

$$E(Y_i(1)|T_i = 1, X_i) - E(Y_i(0)|T_i = 1, X_i) = E(Y_i(1)|T_i = 1, e_i(X_i)) - E(Y_i(0)|T_i = 0, e_i(X_i)).$$

In other words, once we have found countries that are equally likely to be majoritarian conditional on X, a comparison of SMD and PR countries provides a less biased estimate of the causal effect of SMD on SMD countries.

Specifically, King and Zeng (2002) discusses the sources of bias in estimating causal effects, generalizing Heckman, Ichimura and Todd (1997). Bias in causal inference is due to four sources: (a) omitted variable bias, (b) controlling for variables affected by the main causal variable, (c) non-overlapping bias, and (d) density difference bias.

Bias due to (a) is already well-known: omitting a variable that affects price levels and is correlated with the electoral system may bias the causal estimate. This is most evident in the RK Models 1.1 and 1.3, which is why this paper focuses on the fully-specified Model 1.4, though results hold across all models.

Post-treatment bias (b) occurs when a variable that is an effect of the cause is included as a control. Earlier studies on the effects of electoral systems on government expenditures, welfare spending, income equality, and stability might certainly support such a critique of RK. For example, if SMD affects GDP per capita (e.g., through any of the prior effects) we may not be able to isolate the effect due to SMD systems. Rogowski and Kayser 2002, p. 537 generally recognizes that "[o]f course the choice of electoral institutions is itself ultimately endogenous," yet since their empirical models nonetheless assume exogeneity of these covariates, I do not relax that assumption here.[7]

Matching concerns itself with extrapolation due to *non-overlapping bias* (c), stemming from the fact that SMD countries take on certain values of X which no PR countries do, and *density difference bias* (d), stemming for example from the fact that while both SMD and PR countries have low levels of arable land, many more PR observations take on low values than SMD observations. These multidimensional difference are succinctly summarized by density graphs of the propensity scores by electoral system. In such graphs, ranges of where the density of the propensity score between SMD and PR countries differ indicate bias due to (d), whereas a nonzero density of SMD (PR) and a zero density of PR (SMD) would indicate bias due to (c).

Note also that this matching technique is semi-parametric. That is, at this point we make no functional form assumptions about effect of X on Y and we only care to get a balance of X, thereby addressing problems arising from extrapolation. In fact, this method makes uniformly fewer assumptions that regression analysis.[8] The key simply lies in obtaining good matches, or more specifically, in obtaining matching treatment and control groups that have comparable propensity scores. The procedure is akin to matching in the earlier case of one covariate. We match SMD and PR countries that have similar propensity scores, and calculate the difference in means of the outcome variable to estimate the treatment effect.

4.1 Findings

The left panel in Figure 4 graphs the estimated propensity scores for all countries by electoral system, using the covariates from RK Model 1.4. The propensity scores are estimated using a logit model, where $P(T_i = 1|\beta, X) = \frac{1}{1+e^{-x_i\beta}}$. Since the primary role of the propensity score is to balance covariates, the actual

[7]If anything, allowing for cross-cutting effects would yield more difficulties identifying the causal effect. Hence, the standard errors here are, if anything, likely to be conservative.

[8]More formally, propensity score matching assumes that $(Y(1), Y(0)) \perp\!\!\!\perp T|X$ and $0 < Pr(T = 1|X) < 1$, where independence $\perp\!\!\!\perp$ indicates that $P(T, Y(1), Y(0)|X) = P(T|X)$. These are akin to the exogeneity assumption in regression analysis, namely that treatment is assigned based on observable differences (i.e., no omitted variable bias) and of course that X covariates are not a result of treatment, but without functional form assumptions. In order for regression analysis to converge to the matching estimates one must additionally assume (1) constant additive treatment effect, $Y_i(1) - Y_i(0) = \alpha \forall i$, and (2) linearity in all covariates, $E(Y_i(0)|X_i) = X'\gamma$.

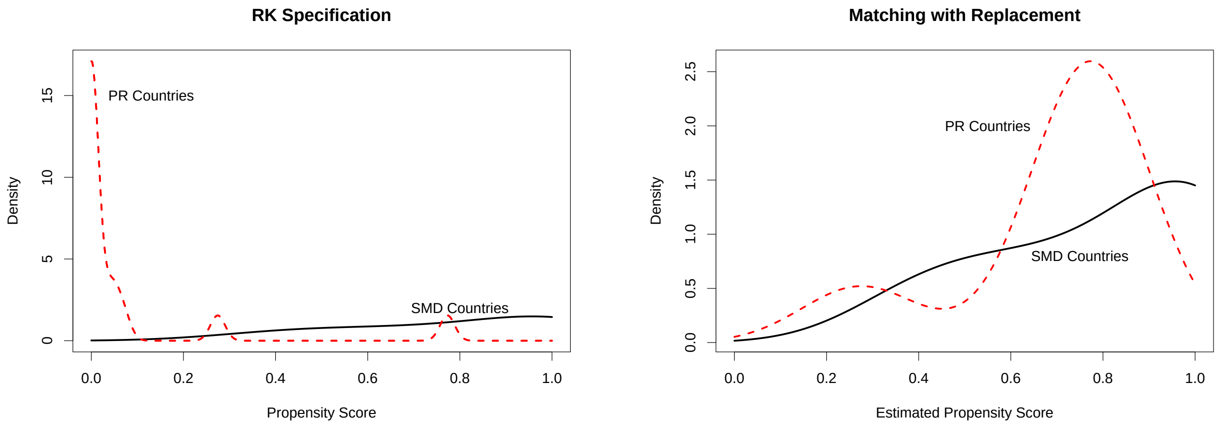

Figure 4: Density estimates (smooth versions of histograms) of estimated propensity scores for SMD and PR countries in original specification of Model 1.4 (left panel) and matching model (right panel).

β's in this step of the estimation are of little substantive significance. I present findings only for Model 1.4 because this is the fully specified model that is least likely to suffer from omitted variable bias, though similar results hold for the other specifications.

The panel summarizes the drastic differences we uncovered before. In fact only one PR country even falls within the range of SMD country propensities (Spain, $\hat{e}_i(X) = 0.773$)! Given this drastic difference, parametric estimates using this data are subject to non-overlapping and density difference bias.

The right panel of Figure 4 presents the distributions of SMD countries after having matched comparable PR countries using a nearest-neighbor matching algorithm with replacement.[9] While the distributions are still discernably different, some common support is at least obtained.

Recall though that the only purpose of the propensity score is merely to balance the pre-treatment covariates. Table 4 presents mean differences between SMD countries and matched PR countries after matching. Matching substantially reduces differences of SMD and PR countries arable land, energy, and import penetration.[10]

Table 5 summarizes the estimates of the treatment effects. Given that matching with replacement is preferable when treatment and control units are vastly different (see Dehejia and Wahba 2002), our best estimate yield a *highly uncertain* treatment effect if it exists at all: after accounting for systematic differences

[9]The procedure for nearest-neighbor matching, the simplest of all matching algorithms, is as follows. Given X_i, I estimate $e_i(X_i) = Pr(T_i = 1|X_i)$. For an SMD country i, I then select some PR country j so as to minimize the absolute distance between propensity scores, $j(i) = \text{argmin}_j |(e_i|T_i = 1) - (e_j|T_j = 0)| \forall j \neq i$. Each SMD country is thereby matched with a PR country. To estimate the average treatment effect, we simply subtract outcomes of the matched SMD and PR countries. Depending on the bias-variance tradeoff in the data, other aspects of matching, such as whether to discard non-overlapping units or whether to match with replacement, become relevant. For a more general reference on matching algorithms, see Rosenbaum (2nd edition, 2002).

[10]Despite the fact that import penetration and size are specified in the Rogowski-Kayser models, they are component parts of trade openness and factor endowment instruments and still balance with inclusion of Model 1.4 covariates in the balancing score.

	T Statistics	
Variable	RK Model	Matching Model
Per Capita GDP	0.23	0.00
Trade Openness	-1.67	-0.98
3 Year Exchange Rate Appreciation	-0.98	-0.89
Log of Arable Land /Population	-2.18^{*}	-1.98
Log of Population	-1.29	-0.71
Log of Energy Production /Consumption	-2.44^{*}	-0.52
Imports as a proportion of GDP	2.93^{*}	1.58
Size (1,000 km2)	-2.26^{*}	-2.18^{*}

Table 4: T-statistics comparing single-member district countries (SMD) and matched proportional representation countries (SMD) on selected covariates. * signifies $|p| < 0.10$. $N = 22$.

Method	Effect of SMD, ATE	Std Err
Matching - No Replacement	-11.3	12.5
Matching - Replacement	-15.6	15.7
RK Estimate	-10.5	5.0

Table 5: Estimated average treatment effects with one-to-one nearest neighbor matching, without discarding. All matching models match six SMD countries with six PR countries. Standard errors were calculated using 500 bootstrapped samples.

between SMD and PR countries, majoritarian systems lead to a 16% decrease in prices, plus or minus 31%! The standard errors are more than three times as high as in the RK findings.

Matching on the propensity score exposes the small sample size problem in the RK data: only six matched PR countries remain after we discount incomparable PR countries, leaving us with highly uncertain estimates. Hence the only way to reassess the price-level effect is through a larger dataset.

5 A Test with New Data on all Democracies

As a preliminary matter, note that RK offers no theoretical reason to believe *a priori* that the price-level effect should be restricted only to OECD countries. Indeed if the RK argument is correct, we should observe this dynamic across all democratic countries. I thereby assemble a new dataset using the same sources and specifications of the original RK study where possible to investigate the price-level effect across all democratic countries, hopefully allowing us also to obtain better matches. Data was gathered for all 73 democratic countries in Perrson and Tabellini (2003), who based their selection of democracies on the Gastil and Polity indices. Data description and sources are provided in Table 9 in the Appendix. The

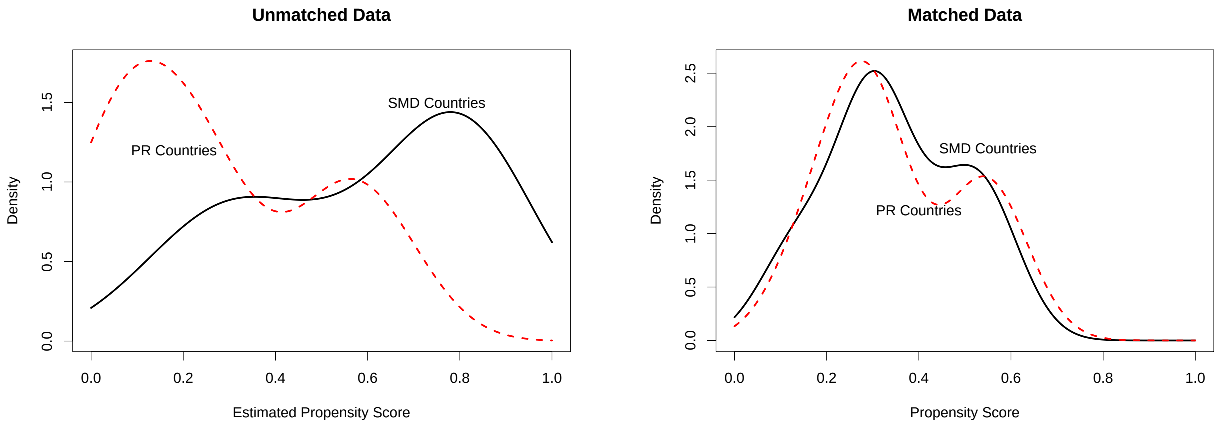

Figure 5: Density estimates of estimated propensity scores for the new dataset without matching (left panel) and with matching (right panel).

only data source that differs from RK consists of the World Development Indicators for domestic energy production, which in the original RK source is only available for OECD countries. Combining these data sources changes none of the conclusions herein. The new data exhibits a substantial missing data problem, with so few observations that make even multiple imputation infeasible. Nevertheless, the new data allows to potentially improve the comparable matched countries to obtain a less biased estimate.

5.1 Matching

With an expanded dataset and new observations, we observe a substantial reduction in the extrapolation problem. The left panel in Figure 5 graphs the density estimates of PR and SMD countries, showing that there is significantly more overlap in the support (non-zero density).

Table 6 provides means tests for SMD and PR countries. SMD and PR countries appear to exhibit fewer systematic differences in this new dataset, largely reflecting the high variance in most covariates with the inclusion of the more heterogeneous non-OECD countries. For example, no longer are all the SMD countries in the dataset large and rich in factor endowments, as evidenced by the addition of SMD countries such as Malaysia and Malawi. The assumption that we can match well is therefore likely to hold with this new data.

After performing nearest-neighbor matching, the balance is substantially improved. The right column of Table 6 summarizes the balance of covariates across all matched SMD and PR countries. There are no statistically significant differences between the matched PR and SMD countries, and all means differences except for those of population decrease with matching.

Most importantly, as this diagnostic already indicates, we now obtain better matches by matching with

Variable	SMD Mean	PR Mean	T-Statistic	Matched
GDP per capita	6.10	10.54	1.87	−0.04
Trade Openness	0.21	0.25	2.06*	−0.18
Exchange Rate	−53.91	−127.00	−0.88	−0.64
Arable Land	0.18	0.23	0.98	−0.18
Population	9.66	9.19	−1.10	1.42
Energy	0.95	0.72	−1.01	−0.73
Imports	38.46	43.14	0.79	0.77
Size (1,000 km2)	340	1011	1.69	1.05

Table 6: T-statistics of new dataset comparing single-member district countries (SMD) and matched proportional representation countries (PR) on selected covariates. $N = 67$.

replacement and discarding non-overlapping units.[11] In the original dataset, discarding non-overlapping units would have left only two total countries. The right panel Figure 5 depicts the density estimates of the propensity scores for matched countries. Note that the densities are close – a drastic improvement over the matches using the original data. This being the crucial assumption of matching methods, estimating the causal effect now becomes a simple difference in means calculation.[12]

Table 7 presents the estimates of the causal effect of a majoritarian electoral system on price levels. We obtain the best matches by discarding outliers and matching with replacement, thereby minimizing non-overlapping and density difference bias. This yields a treatment effect SMD on prices of roughly 2.1%, plus or minus 35% – again, a highly uncertain effect. From the left columns we can see that outliers appear to be driving the price-level effect, and even then not at any significant levels. In fact, with these specification, the price-level effect confidence interval ranges from as high as 37% to as low as -60% in the matching models. Again, we discover the bias-variance tradeoff: the RK price-level effect is an artifact of the functional form imposed by the regression, and by comparing similar PR and SMD countries (using the same exact datasources and covariates), we find little evidence for the broad price-level effect. Figure 6 graphs the substantive difference between the RK and the matching models, demonstrating the tradeoff.

Lastly, note it is possible that the new data introduces more *unobserved* heterogeneity by including non-OECD countries. Capturing these differences between OECD and non-OECD countries by matching on an indicator variable for OECD, however, yields the same inconclusive price-level effect. Nevertheless, to the

[11] Specifically, treatment units with propensity scores above the maximum propensity score of control units and control units with propensity scores below the minimum of propensity scores of treatment units were discarded. This addresses non-overlapping bias. As King and Zeng point out, these discarded units in themselves may be of substantive interest, allowing us to determine which counterfactuals are reasonable comparisons, akin to choosing case studies. Replacement reduces density difference bias by providing better matches when treatment and control distributions are quite different.

[12] Estimation of the standard errors is not quite this simple, requiring us to bootstrap the entire sample, which is standard in the literature (see, e.g., Dehejia and Wahba (2002), Imai (2003)). The bootstrap procedure draws samples from the donor pool data with replacement, matches on the propensity score within each bootstrapped sample, then calculates the ATE for the matched units in the bootstrapped sample, and finally obtains the variance across the bootstrapped estimands.

	Discard Outliers					
	(1) No			(2) Yes		
Matching Method	ATE	SE	Matches	ATE	SE	Matches
No Replacement	-17.6	9.0	13	-0.1	14.9	6
Replacement	-19.0	20.3	13	2.1	17.5	6
RK Specification				-10.5	5.0	

Table 7: Estimated average treatment effects (ATE) of majoritarian electoral systems on prices with one-to-one nearest-neighbor matching. Standard Errors (SE) calculated using 500 bootstrapped samples. Matches refers to number of paired SMD and PR countries.

degree that there are important omitted variables, a criticism that would just as well apply to the original dataset, it would be productive for researchers to gather this additional data.[13]

6 Conclusion

These results significantly question the "clear finding that . . . prices of goods and services are systematically higher in PR countries" (Rogowski and Kayser 2002, p. 526). Several comments bear note here. First, these findings do not necessarily undermine the important line of research of RK. Much to the contrary, this paper suggests locating the causal mechanism with a finer comb. Perhaps a more refined measurement of seats-votes elasticity (which could be district-specific) might pinpoint the price-level effect more specifically (see, e.g., King 1989). Perhaps the assumption that majoritarian systems have categorically higher seats-votes elasticities does not fit the empirical reality of multiple parties and vote shares that might deviate widely from a 50% vote share. Recall that the two-party and roughly 50% vote share assumptions were crucial to this claim, which would suggest stark omitted variable bias due to exclusion of expected voteshares. Or perhaps prices are too rough an indicator of the tradeoff between consumers and producers. What is certain, however, is that the systematic differences between SMD and PR countries make a convincing empirical test of the price-level effect much more difficult than originally claimed. Nonetheless, any of these above refinements might help us pinpoint more accurately where and under what conditions the causal effect might exist. Rather than discretize electoral rules into SMD and PR systems we may need to examine the complexities of electoral rules that Cox (1990) for example has classified. Changes in the electoral system in Japan, New Zealand, the Philippines, and the Ukraine in the 1990s might also prove fruitful in narrowing down the effect of electoral systems on price levels. Such "natural experiments" arguably adhere more closely to the assumptions of randomized treatment assignment (see, e.g., Rosenbluth and Schaap

[13] It is of course possible to conduct a sensitivity analysis here due to omitted variables, as outlined in Rosenbaum and Rubin (1983a) and Imbens (2002), but such an analysis would likely serve only to widen confidence intervals. In other words, assuming exogeneity, the standard errors here are if anything *conservative*.

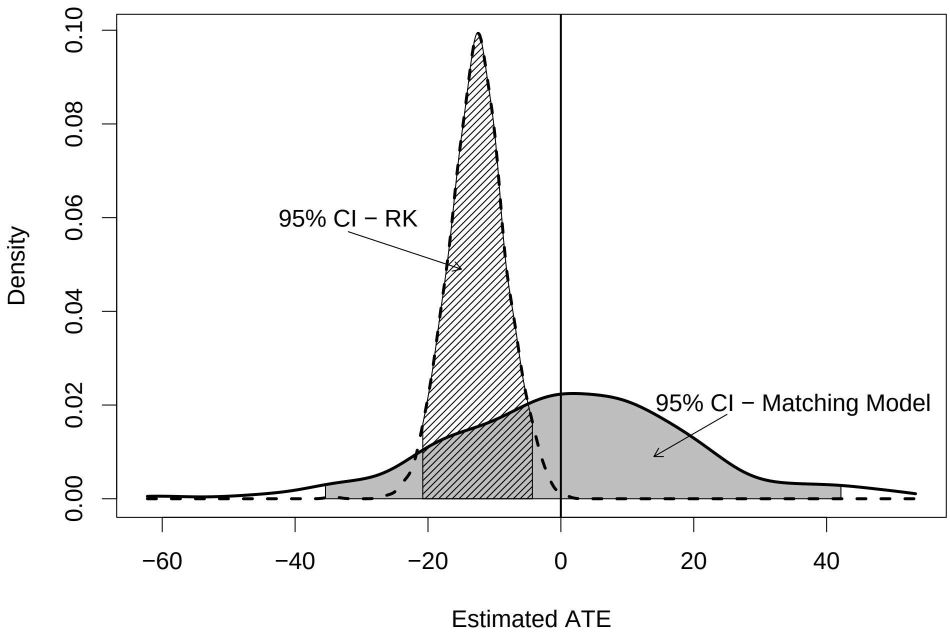

Figure 6: Bias-variance tradeoff: Comparison of average treatment effect of majoritarian system on price levels, simulated from RK Model 1.3 (holding other variables at median values) and matching model with new dataset (with replacement and discarding outliers). Shaded areas represent 95% confidence intervals, and vertical line represents no impact of SMD on price levels.

2002). Second, the methods developed should prove useful not only to quantitative researchers seeking to determine what their data can tell them, but also to qualitatively-oriented researchers seeking to formalize case selection criteria. Propensity score matching permits researchers to choose good cases, thereby aiding in the design of their study before even venturing into the field. Lastly, one could well retreat from the position that electoral systems have a *causal effect* on prices, but that arguably requires abandoning one, if not the, central claim of RK.

A Appendix

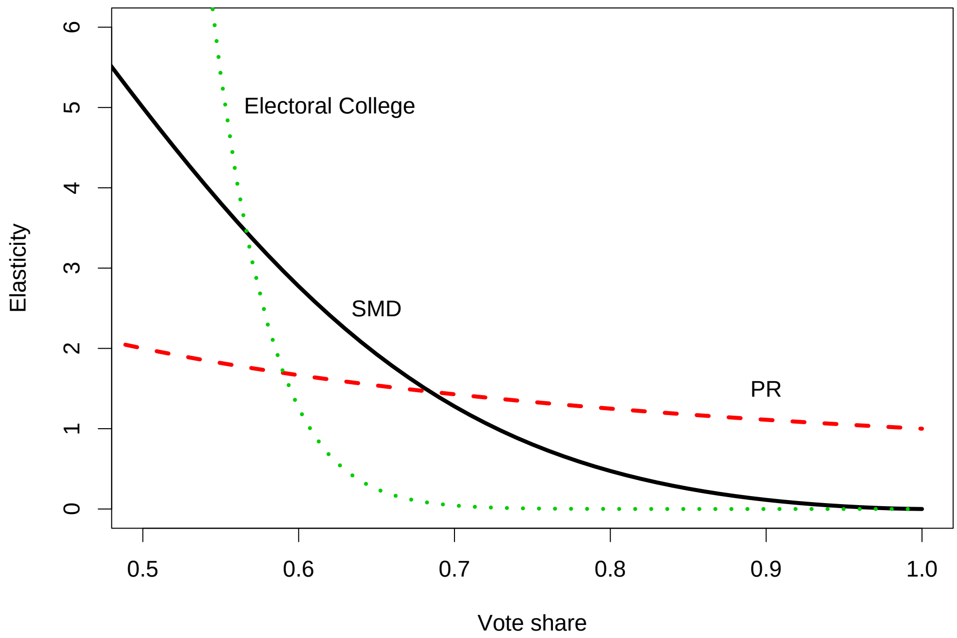

Figure 7: Seats-Vote elasticity as a function of vote share in three electoral systems. Replicated from RK Note 15, p. 531.

	Dependent Variable: GDP Price		
Constant	88.37***	100.97***	90.84***
	(16.41)	(19.61)	(16.48)
GDP per capita	2.81***	.	2.82***
	(0.29)	(0.49)	(0.29)
Majoritarian	**-6.93**	**-10.68**	**-9.88**
	(14.35)	**(6.58)**	**(6.04)**
Trade openness	-68.00**	-82.61**	-68.88**
	(23.06)	(28.89)	(23.01)
Exchange Rate Appreciation	0.12**	0.22**	0.12**
	(0.04)	(0.05)	(0.04)
Ln(Arable land/pop)	3.54	4.88	1.90
	(10.63)	(9.86)	(7.89)
Ln(Population)	-2.58	-1.30	-2.70
	(1.53)	(1.89)	(1.49)
Ln(Energy)	-3.08	5.42	-0.42
	(4.63)	(5.60)	(18.08)
Maj * Ln(Energy)	**-5.23**	.	.
	(19.88)	.	.
GDP^2	.	0.06***	.
	.	0.01	.
$Ln(Energy)^2$	.	.	1.40
	.	.	(7.80)
R^2	0.97	0.94	0.97
F	45.97	31.1	45.84
N	22.	22.	22.

Table 8: Implications of extrapolation – effects of interaction and squared terms on original estimates. Bolded figures indicate substantive impact of SMD on price levels. Standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Variable	Description	Source
Price Level	Price level GDP [%] (PPP GDP/ U.S. dollar exchange rate)	Penn World Tables, 5.6
Majoritarian electoral system	Dummy variable scored as 1 if all the lower house is elected under plurality rule, 0 otherwise	Perrson and Tabellini (2003)
GDP per capita	GDP per capita in US$1,000	GDP: IMF IFS rf..zf, Population: Penn World Tables, 5.6
Exchange Rate Appreciation	Percentage change in NC/USD exchange rate since 1987	IMF IFS rf..zf
Arable land	Log of ((arable land in hectares / population)+1)	World Development Indicators (2002)
Energy	Log of (commercial energy production (kt of oil equivalent) / commercial energy use (kt of oil equivalent))	World Development Indicators (2002)
Population	Natural log of population in millian inhabitants	Penn World Tables, 5.6
Openness	Measure of "Free trade openness" instrumented as $Openness = 0.528 - 0.026 log(area) - 0.095 log(dist)$, where *area* represents the size of the country in million square km and *dist* represents the average distance to capitals of world 20 major exporters, weighted by values of biliateral imports in 1000 km	Lee (June 1993)

Table 9: Variable description and data sources of new dataset including non-OECD democracies

References

Cox, Gary. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34:903–35.

Dehejia, Rajeev H. and Sadek Wahba. 1999. "Causal Effects in Nonexperimental Studies: Re-Evaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94:1053–62.

Dehejia, Rajeev H. and Sadek Wahba. 2002. "Propensity Score Matching Methods for Non-Experimental Causal Studies." *Columbia University, Department of Economics Discussion Paper* 0102-14.

Fiona, McGillivray. 2002. "Government Hand-Outs, Political Institutions, and Stock Price Dispersion." Manuscript, available at `http://www.nyu.edu/gsas/dept/politics/seminars/fiona.pdf`.

Gourevitch, Peter A. 2003. "The Politics of Corporate Governance Regulation: Political Determinants of Corporate Governance: Political Context, Corporate Impact (book review)." *Yale Law Journal* 112:1829–80.

Gourevitch, Peter, Richard Carney and Michael Hawes. 2003. "Testing Political Explanations of Corporate Governance Patterns." Manuscript, available at `http://www.stanford.edu/group/sshi/Finance2003/gourevitch_paper.pdf`.

Heckman, James J., Hidehiko Ichimura and Petra Todd. 1997. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 64:605–54.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945–60.

Imai, Kosuke. 2003. "The Importance of Statistical Methodology for Analyzing Data from Field Experimentation: Evaluating Voter Mobilization Strategies." Manuscript, presented at 2002 Annual Meeting of the American Political Science Association, Boston, MA, and available at `http://www.people.fas.harvard.edu/~kimai`.

Imbens, Guido W. 2002. "Sensivity to Exogeneity Assumptions in Program Evaluation." Manuscript, available at `http://emlab.berkeley.edu/users/imbens/imbens.pdf`.

Iversen, Torben and David. Soskice. 2002. "Insurance and Representation: Why do Some Democracies Redistribute More than Others?" Manuscript, presented at 2002 Annual Meeting of the American Political Science Association, Boston, MA.

King, Gary. 1989. "Representation through Legislative Redistricting: A Stochastic Model." *American Journal of Political Science* 33:787–824.

King, Gary and Langche Zeng. 2002. "When Can History Be Our Guide? The Pitfalls of Counterfactual Inference." *Manuscript, available at* `http://Gking.harvard.edu`.

King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Lee, Jong Wha. June 1993. "International Trade, Distortions and Long-run Economic Growth." *IMF Staff Papers*.

Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Milner, Helen V. and Benjamin Judkins. 2002. "Partisanship and Trade Policy: Is there a Left-Right Divide on Trade Policy?" Manuscript, presented at 2002 Annual Meeting of the American Political Science Association, Boston, MA.

Pablo T. Spiller, Ernesto Stein and Mariano Tommasi. 2003. "Political Institutions, Policymaking Processes, and Policy Outcomes: An Intertemporal Transactions Framework." Manuscript, available at `http://www.wws.princeton.edu/~rppe/apr2003conf/Tommasi.doc`.

Perrson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Cambridge, MA: MIT Press, Forthcoming.

Rogowski, Ronald, Eric C. C. Chang and Mark Andreas Kayser. 2002. "Electoral Systems and Real Prices: Panel Evidence for the OECD Countries, 1970-2000." Manuscript presented at 2002 Annual Meeting of the American Political Science Association, Boston, MA.

Rogowski, Ronald and Mark Andreas Kayser. 2002. "Majoritarian Electoral Systems and Consumer Power: Price Level Evidence from the OECD Countries." *American Journal of Political Science* 46:526–39.

Rosenbaum, Paul R. 2nd edition, 2002. *Observational Studies*. New York, NY: Springer Verlag.

Rosenbaum, Paul R. and Donald B. Rubin. 1983a. "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome." *Journal of the Royal Statistical Society, Series B* 45:212–18.

Rosenbaum, Paul R. and Donald B. Rubin. 1983b. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

Rosenbluth, Frances and Ross Schaap. 2002. "The Domestic Politics of Banking Regulation." Manuscript, presented at 2002 Annual Meeting of the American Political Science Association, Boston, MA.

Scartascini, Carlos G. 2002. "Political Institutions and Regulation: An Inquiry on the Impact of Electoral Systems on the Regulation of Entry." Manuscript, available at `http://www.gmu.edu/departments/economics/bcaplan/scar.pdf`.

Scartascini, Carlos G. and W. Mark Crain. 2002. "The Size and Composition of Government Spending in Multi-Party Systems." Manuscript.