Democracy and Human Development

John Gerring Boston University
Strom C. Thacker Boston University
Rodrigo Alfaro BBVA Research

Does democracy improve the quality of life for its citizens? Scholars have long assumed that it does, but recent research has called this orthodoxy into question. This article reviews this body of work, develops a series of causal pathways through which democracy might improve social welfare, and tests two hypotheses: (a) that a country's level of democracy in a given year affects its level of human development and (b) that its stock of democracy over the past century affects its level of human development. Using infant mortality rates as a core measure of human development, we conduct a series of time-series—cross-national statistical tests of these two hypotheses. We find only slight evidence for the first proposition, but substantial support for the second. Thus, we argue that the best way to think about the relationship between democracy and development is as a time-dependent, historical phenomenon.

From classical Greece to the present era, writers have usually assumed that the institutions of democracy propel a political dynamic favorable to the needs and interests of the less advantaged citizens in a society.[1] While there has always been controversy over the question of whether democracy enhances economic development (Kurzman, Werum, and Burkhart 2002; Przeworski et al. 2000), the consensus view has been that democracy enhances human development (e.g., Boix 2001; Brown and Hunter 2004; Brown and Mobarak 2009; Dreze and Sen 1989; Ghobarah, Huth, and Russett 2004; Kudamatsu 2006; Lake and Baum 2001; Lenski 1966; Lipset 1959; Muller 1988). The logic of this argument rests largely on the idea that popular participation in government empowers ordinary citizens—including the very poor—and should, as a result, lead governments to be more accountable to their interests. For Aristotle, Madison, and most latter-day political economists, it is nearly axiomatic that democracy serves as a mechanism for redistribution (Meltzer and Richard 1981).

Recently, this consensus opinion has been strongly challenged. Several studies argue that there is no positive correlation between regime type and various measures of human development or that these relationships are not especially robust (Gauri and Khaleghian 2002; McGuire 2004; Ross 2006; Shandra et al. 2004). These large-sample cross-national studies are bolstered by a good deal of qualitative evidence. Some of the most dramatic improvements in human development over the course of the twentieth century have occurred under the auspices of authoritarian rule (e.g., in the East Asian NICs and in communist countries), while many democratic societies in the developing world have been characterized by persistent disparities in wealth and high levels of poverty (e.g., India, sub-Saharan Africa, and many Latin American countries). Moreover, some of the causal pathways by which democracy was previously thought to influence the welfare of the poor seem rather dubious in light of recent empirical analysis. While conventional wisdom assumed (largely on the basis of the experience of OECD countries) that democracy would lead to higher social spending, and this, in turn, would enhance the welfare of the poor, it turns out that there is little or no correlation between public spending and human development outside the OECD (Filmer and Pritchett 1999; McGuire 2004). The stipulated mechanisms of the welfare state do not lead—at least not in any consistent fashion—to an

[1]An online appendix with supplementary material is available at http://journals.cambridge.org/jop. Data and supporting materials necessary to reproduce the numerical results will be made available at www.bu.edu/sthacker no later than January 2012.

The Journal of Politics, Vol. 74, No. 1, January 2012, Pp. 1–17 doi:10.1017/S0022381611001113

ISSN 0022-3816

improvement in social welfare as measured by mortality, literacy, and other human development outcomes. Thus, even if one brackets the question of economic growth (thereby assuming that regime type is growth neutral), the case for democracy as a welfare-enhancing mechanism appears shaky. An age-old assumption faces serious challenge.

Thus far, the debate between proponents and skeptics has centered largely on democracy's contemporaneous relationship to human development. Empirical work tests the relationship between democracy today and human development in the following year or decade (depending upon the time lag of the model). Theoretical discussions follow suit. In this article, we introduce the possibility that democracy's developmental effects might be longer-term, characterized by a distal rather than proximal causal relationship.

We begin by revisiting the traditional theoretical arguments in an attempt to show what a difference history might make. We proceed to test two hypotheses in a series of cross-national regression tests with the infant mortality rate (IMR) as our primary measure of human development. The first hypothesis replicates the traditional causal model, linking IMR to a country's *level* of democracy in the previous year. The second hypothesis measures democracy with a *stock* index that captures a country's regime history from 1900 to the observation year. Our findings demonstrate that a country's contemporary level of democracy has only a weak association with improved human development while a country's historical experience with democracy has a strong and robust influence on human development. We conclude that democracy advances human development, but only when considered as a historical (''stock'') phenomenon.

Thinking Historically about Democracy and Development

For the most part, the two sides of the longstanding debate between those who are optimistic about democracy's effect on human development and those who are more skeptical share one fundamental presupposition: democracy has a proximal relationship to developmental outcomes. Yet, new democracies and old democracies are not the same. While new democracies are prone to a host of problems associated with regime transition, older, more institutionalized democracies generally enjoy higher-quality governance (Kapstein and Converse 2008; Keefer 2006). It would be surprising, indeed, if the human development performance of countries moving from authoritarian to democratic rule were substantially improved over the course of the subsequent year or decade. We surmise, however, that if a democratic form of government is maintained over a longer period of time the net effect of that regime type will be positive for the welfare of its citizens.

Note that regimes do not begin again, de novo, with each calendar year. Where one is today depends critically upon where one has been. Historical work suggests that democracy and authoritarianism construct deep legacies, extending back several decades, perhaps even centuries (Collier and Collier 1991; Hite and Cesarini 2004). It follows that we should concern ourselves with the accumulated effect of these historical legacies, not merely their contemporary status. We contend, therefore, that the effects of political institutions are likely to unfold over time—sometimes a great deal of time—and that these temporal effects are cumulative. Let us consider four of the numerous possible causal pathways linking democracy and human development (McGuire 2004; Ross 2006), taking into account the possible time-dependent nature of this relationship.

First, competition among elites for voters' favor should produce a situation in which elites are accountable to the citizenry—or, at the very least, to a plurality of the voting electorate. Since widespread human misery is unpopular, democratically elected leaders may be more likely to concern themselves with issues of human development than leaders who maintain their positions through other means (Lake and Baum 2001). To be sure, authoritarian leaders might also be concerned with the potentially destabilizing effects of widespread poverty. However, they may be more likely to weather this kind of bad news than their democratic counterparts because they face a much smaller selectorate (Bueno de Mesquita et al. 2003). As long as the authoritarian regime's core constituency (e.g., the military, ruling party, and economic elites) is well compensated, it is unlikely that the sufferings of the masses will threaten their control over the state. A striking example of this can be found during the China's Great Leap Forward (Kane 1989; Riskin 1995). The massive starvation that ensued in the wake of Mao's disastrous reforms, which may qualify as the largest number of government-induced deaths in recorded history, did not threaten Mao's leadership or the leadership of the Chinese Communist Party. It is difficult to imagine such an event occurring in a democratic setting

without serious negative consequences for those in power.

The accountability argument is quite plausible when applied to disastrous policy outcomes such as famine, and the empirical results are strong. To date, no large-scale famine has occurred in a full-fledged democracy (Dreze and Sen 1989). Yet, for more complex developmental policies, where the failures are less obvious, less dramatic, and less easily tied to the current government, the principal-agent logic attenuates. There is no obvious reason why a democratically elected government would benefit from incurring present costs for the sake of future gains unless the time horizons of those elites have shifted to a longer-term perspective. Indeed, we do not expect long-sighted policies to emanate from a recently democratized polity, where institutions are in flux, parties are nascent, and voter affiliations ephemeral. Faced with political uncertainty and instability, politicians face incentives to pursue short-run goals at the expense of long-term development (Haggard 1991; Keefer 2006). Indeed, in a new democracy there is little assurance that the democratic framework will hold; a country's most recent election may be its last. Under these circumstances, it is understandable that politicians and voters might adopt policies where the pay-offs are short-term rather than long-term. In a longstanding democracy, by contrast, it seems plausible that leaders might find it in their interest to pursue policies where the benefits lie far off in the future. Here, actors may assume that free and fair elections will continue. Regular alternation in power lowers the short-run stakes of a given election. The losing party may reasonably expect to regain power at a later date. This means that meritorious actions taken while a party is in office may have beneficial consequences for that party long after their term of office has expired. As an example, one might consider the enormous political benefits reaped by the (U.S.) Democratic Party in the postwar era from the passage of the Social Security Act in 1935.

Second, the institutions of democracy tend to foster a well-developed civil society. This is because political rights and civil rights are highly correlated, and the existence of civil rights usually leads, over time, to a dense network of voluntary associations, which may be religious or secular, national or international, issue-specific or broadly pitched (Parker 1994). In turn, these voluntary associations are often instrumental in providing services for the poor, perhaps in conjunction with official state bodies and/or international actors. They may also be instrumental in lobbying for legislation that addresses the needs of the poor and improves the quality of public administration (Sondhi 2000). Nongovernmental organizations (NGOs) appear to have played a critical role in child vaccination campaigns, in campaigns for the treatment of HIV/AIDS, in education and health care, and in many other policies that directly affect the general welfare (Gauri and Khaleghian 2002; Gauri and Lieberman 2006; Lake and Baum 2001; McGuire 2010). The evolution of civil society is a long-term process. Voluntary associations and NGOs do not spring forth overnight. Thus, insofar as strong civil societies encourage better governance and greater attention to the needs of the less advantaged citizens in a society, we can expect these causal mechanisms to kick in only with the passage of time. Again, the age of democracy would seem to matter when considering human development outcomes.

Third, democracy may serve to inaugurate a culture of equality that empowers oppressed groups. In the process of granting formal citizenship rights to outgroups—lower castes and classes, peasants, racial, ethnic, and religious minorities—democracy may foster a political dynamic in which these groups conceptualize their interests as a matter of rights and take a correspondingly aggressive approach to satisfying those rights in the political, social, and economic spheres (Alvarez, Dagnino, and Escobar 1998; Piven and Cloward 1977; Rubin 1997). This political dynamic, once initiated, may have important repercussions for societal human development insofar as it leads to an extension and improvement of government services and increased utilization of those services. Again, it seems reasonable to suppose that this process of change would occur only over a period of decades, if not centuries (e.g., the American civil rights movement).

Finally, we expect that older democracies will benefit from greater institutionalization in the political sphere. Although political institutionalization is difficult to define, there seems to be general consensus that procedures in a well-institutionalized polity are functionally differentiated, regularized (and hence predictable), professionalized (including meritocratic methods of recruitment and promotion), rationalized (explicable, rule based, and nonarbitrary), and infused with value (Huntington 1968; Levitsky 1998; Polsby 1968). Most long-standing democracies fit this description. They feature highly developed, highly differentiated systems of governance, involving both formal bureaucracies and extra-constitutional organizations such as interest groups, political parties, and other nongovernmental organizations. Arguably, the length of time a democracy has been in existence

serves as a rough indicator of its degree of institutionalization. By contrast, the length of time an authoritarian regime has been in existence may have little or no bearing on its level of institutionalization. Witness: the latter days of Nicaragua under Somoza or Iraq under Saddam Hussein.[2]

We suspect that the reasons for this stem directly from their systems of rule. Where power is personalized, as it is in many authoritarian settings, the development of legal-bureaucratic authority is virtually impossible. In particular, leadership succession is difficult to contain within regularized procedures and promises a period of transition fraught with uncertainties. Thus, even if a monarch or dictator adheres to consistent policy objectives during his or her rule, there may be little continuity between that regime (a term we employ here in its broader sense) and its successor. The hallmark of a long-standing democracy, by contrast, is its ability to resolve the problem of leadership succession without turmoil and without extraordinary discontinuities in policy and in political organization. The framework remains intact, and this means that the process of institutionalization may continue, despite the occasional bump in the road.

More importantly, we suspect that the institutionalization of power leads to greater gains within a democratic setting than in an authoritarian setting. Institutionalization matters more under democracy. Consider the problem of establishing social order and stability in a polity and resolving problems of coordination (Hardin 1999). Noninstitutionalized polities are unstable and inefficient, almost by definition, for there are no regularized procedures for reaching decisions. However, in an authoritarian setting, a Hobbesian order may be established simply and efficiently by fiat and force. Rule by coercion, insofar as it is successful, can be imposed with scant loss of time and little negotiation; the threat of force is immediate. Consequently, there is less need for highly institutionalized procedures for reconciling differences and establishing the force of law. The sovereign may rule directly.

In a democratic setting, by contrast, resolving conflict is complicated and generally takes a good deal of time. Somehow, everyone must agree upon (or at least agree to respect) the imposition of society-wide policy solutions that involve uneven costs and benefits. In order to handle these quintessentially political problems, a democratic polity has little choice but to institutionalize procedures for negotiation among rival constituencies and organizations. Once these procedures are established, we expect them to be more effective in resolving differences and finding optimal solutions than would be fiats imposed from above. Indeed, whatever centripetal tendencies are inherent in democracy are more likely to be in evidence when those democratic arrangements have been in operation for some time. For this reason, the thesis of democratic overload is much more compelling when applied to new democracies than when applied to old. Democratization is often a boisterous, obstreperous affair. Established democracies, by contrast, tend to be more restrained. In particular, the norm of incremental change is more likely to be accepted.

Thus, given sufficient time, we expect that democracies will provide greater stability and more efficient public policies. Arguably, the problem of overload arises not from institutional sclerosis (Olson 1982) but rather from insufficient institutionalization (Huntington 1968). This provides yet another reason to suppose that long-term democracies may adopt smarter social policies and may implement them with greater perseverance and flexibility. If democracy survives its often tumultuous youth, indicators of human development should demonstrate marked improvements—even if no immediate improvement was registered in the initial transition from authoritarian rule. Democracy, we conclude, is best considered as a *stock*, rather than *level*, concept. Two dimensions of democracy, time and regime type, must be gauged together in order to explain a country's human development capacity.

Measuring Human Development

We understand the concept of human development to refer to the aggregate welfare of societies, with particular attention to less advantaged citizens. As a cross-national and historical measure of human development we rely on the infant mortality rate (IMR), i.e., the number of deaths prior to age one per 1,000 live births (a discussion of this choice and other possible indicators can be found in the online appendix).

Scholars have compiled two important global IMR datasets in recent years, one sponsored by UNICEF (Hill et al. 1999) and the other by the World

[2]Institutional decay can also occur under democratic rule, as some political party systems in Latin America (e.g., Venezuela and Peru) have demonstrated in recent years. Our argument is probabilistic: old democracies are more likely than new democracies and authoritarian regimes to have strong institutions.

Bank (2003; see discussion in Ross 2006). Reassuringly, these two measurements of IMR are nearly perfectly correlated (r=0.996). Not surprisingly, results using either variable are quite similar. We show results only for the latter (drawn from the *World Development Indicators* dataset [World Bank 2003]) because it has broader country coverage and is less vulnerable to sample biases. (The substantive interpretation of the results is the same for both.)

Because we employ a fixed-effects format in most of our statistical tests, the principal methodological issue is less the cross-national comparability of the data than the within-country longitudinal comparability. For example, if surveys of IMR in Sri Lanka employ a somewhat different methodology than surveys of IMR in India this is less problematic than if surveys within Sri Lanka or India change their methodology, without making subsequent corrections in previously collected data. (Some of our analyses also introduce year-specific controls, which should compensate for any global changes in methodology.)

In analyzing the causes of infant mortality it is important to make some correction for the bounded, uneven nature of this phenomenon. Although IMR measures increments of one, we cannot interpret these increments as representing equal policy achievements. (Presumably, it is easier to lower IMR from high levels than from low levels.) Consequently, IMR is usually measured as an elasticity, i.e., as a percentage change in the outcome. We achieve this by transforming infant mortality rates into logarithmic form.

Since IMR data are not available on an annual basis for many countries, we adopt two complementary strategies to avoid the potential bias associated with nonrandom patterns of missing data. First, we interpolate missing data to create a more complete time series for each country. This increases the potential sample from 4213 (as drawn from World Bank 2003) to 7418, a sizeable augmentation. However, in no case do we extrapolate missing data beyond the first and last data points for a country. And in no case are there more than two or three consecutive years of missing data. Note that IMR data are highly regular; in the short run, temporal changes tend to follow well-defined paths for each country. In this situation, the technique of interpolation is relatively unproblematic. (By contrast, one would hesitate to interpolate missing data for growth or inflation, since these factors vary irregularly from year to year.) Given the highly ''predictable'' nature of IMR from year to year, it makes sense to work with a complete annual dataset that is close to what we suspect is the actual trajectory for each country rather than a dataset that is much smaller, irregularly spaced, and likely biased by the systematic omission of data from the poorest countries.

Second, in some analyses we employ the AMELIA technique of multiple imputation (King et al. 2001) to create full samples for analysis. The technique is applied to the original (not interpolated) IMR data, generating a series of imputed datasets with full coverage and increasing the total number of observations to 6,233.[3]

Note that the following analyses are conducted with interpolated IMR, but not additional imputed data, except where indicated. A few analyses are conducted without any data additions (i.e., with the raw IMR data as provided by the World Development Indicators 2003 dataset). Results for key variables are stable across these differently constructed samples.

Measuring Democracy

There is no fully satisfactory measure of regime type (Munck and Verkuilen 2002), and the options are considerably reduced when one requires a measure that provides a large sample of countries over a long period of historical time. We consider it desirable to measure the quality of regimes in as differentiated a fashion is possible, thus precluding dichotomous measures of democracy (e.g., Przeworski et al. 2000). Among the continuous measures, the only measure with broad historical coverage is the Polity2 variable, drawn from the Polity IV dataset (Marshall and Jaggers 2000). This variable measures the extent to which democratic or authoritarian ''authority patterns'' are institutionalized in a given country. It takes into account how the executive is selected, the degree of checks on executive power, and the form of political competition.

The Polity2 variable is, in principle, highly sensitive (coding ranges across a 21-point scale). It also offers extensive country (all sovereign polities except microstates) and historical coverage. Moreover, it allows us to consider both the degree of democracy in any given country-year and its duration over a long period of time (the dataset begins in 1800). The Polity

[3]As recommended by King et al. (2001), the multiple imputation process included all the variables used in this study, plus a series of other variables plausibly related to the variables used here: trade/GDP (World Bank 2003), King and Zeng's (2001) measure of deviation from the global IMR mean, the UNICEF infant mortality measure (Hill et al. 1999), and total illiteracy (World Bank 2003).

dataset, however, imposes two notable costs. First, the rules used to create the key variable, Polity2, are dizzyingly complex. The Polity User's Manual makes a valiant effort to explicate coding procedures, but the methods remain rather difficult to unpack. Second, there are serious questions regarding measurement error in the index (Bollen and Paxton 2000; Bowman, Lehoucq, and Mahoney 2005; Munck and Verkuilen 2002; Treier and Jackman 2008). To be sure, questions might be raised with respect to all extant, and all conceivable, democracy indices. Polity2 is likely no worse than the rest, and probably better than most. It is, indeed, the industry standard, owing largely to the strengths noted above. Reassuringly, it correlates highly with other existing measures of democracy. We do not suspect systematic errors in this index that might affect the substantive findings of this study.

To correct for Polity2's exclusion of microstates, an exclusion that might bias our sample, we impute democracy scores for these excluded cases using other democracy indices that are conceptually and empirically close to the Polity2 measure: (1) the Freedom House Political Rights indicator (see freedomhouse.org), (2) Ken Bollen's Liberal Democracy variable (Bollen 1993), (3) Tatu Vanhanen's Competition variable (Vanhanen 1990), (4) Arthur Banks's Legislative Effectiveness variables (I and II), and (5) Banks's Party Legitimacy variable (Banks 1994). These measures of democracy take into account the degree to which citizens can participate freely in the political process, the extent of suffrage, the competitiveness of national-level elections, the degree of party competitiveness, and the degree to which the legislature affects public policy. With the additional imputed data, the original Polity2 variable gains about 1,500 observations, constituting roughly 20% of the final available sample. (Reassuringly, adding these data points does not have an appreciable effect on the results reported in subsequent tables.)

What we refer to as a *level* measure of democracy is simply the score a country receives on the Polity2 index (scored from -10 to +10) in a given year. To create a *stock* measure of democracy we sum each country's score from 1900 to the present year, applying a 1% annual depreciation rate. This means that a country's regime stock stretches back over the course of the twentieth century but that more distant years receive less weight than recent ones. Our expectation is that the causal effect of democracy, like other capital stocks, depreciates over time. We choose the year 1900 as a threshold period that ushered in a period in which mass democracy becomes a world-historical phenomenon (no longer restricted to the United States and a few European states). We choose a 1% depreciation rate because it seems a reasonable estimation of how a long-run historical effect might play out. (For further discussion of depreciation rates, see below.) An example of how this coding translates into specific ''stocks'' for specific countries is given in the appendix.

Because the historical component of this index weighs heavily on our understanding of the concept and because the Polity dataset ignores nonsovereign states in its coding procedures, we supplement the Polity2 coding with our own coding of several nation-states that were previously part of contiguous empires. The procedure is as follows. For each year that a nation-state belonged to a contiguous imperial power it receives the same Polity2 score as its imperial ruler; e.g., Estonia receives the same score as the Soviet Union from 1941 through 1990. We use this procedure only for nation-states contiguous with the empire to which they previously belonged. We assume that contiguous colonies are likely to be governed in the same manner as the imperial power itself, a dynamic less likely to be true for overseas colonies.[4]

Method of Analysis

The empirical tests consist of a series of cross-national estimations in which we regress the natural log of IMR against democracy, along with various controls. The resulting samples include all countries for which relevant data are available during the 1960 to 2000 time period. This allows for the construction of time series—cross-section samples that approach nearly complete global coverage of sovereign nations

[4]This re-coding affects the following countries: Albania (1900–12, Ottoman Empire), Andorra (1900–present, France), Armenia (1900–90, Russia/USSR), Azerbaijan (Russia/USSR 1900–90), Belarus (Russia/USSR, 1900–90), Bosnia-Herzegovina (1908–17, Austria-Hungary; Yugoslavia 1929–91), Croatia (1900–17, Austria-Hungary; Yugoslavia 1929–91), Czech Republic (1900–17, Austria-Hungary), Slovakia (1900–17, Austria-Hungary), Estonia (1900–16 and 1941–90, Russia/USSR), Finland (1900–16, Russia), Georgia (1900–90, Russia/USSR), Iraq (1900–17, Ottoman Empire), Israel (1900–17, Ottoman Empire), Kazakhstan (1900–90, Russia/USSR), Kyrgyzstan (1900–90, Russia/USSR), Latvia (1900–17 and 1941–90, Russia/USSR), Lithuania (1900–17 and 1941–90, Russia/USSR), Macedonia (1922–90, Yugoslavia), Moldova (1900–45, Romania; 1946–90, USSR), Mongolia (1900–20, China), Bangladesh (1947–71, Pakistan), Slovenia (1900–17, Austria-Hungary; Yugoslavia 1929–91), Syria (1900–17, Ottoman Empire), Tajikistan (Russia/USSR, 1900–90), Turkmenistan (1900–90, Russia/USSR), Ukraine (1900–17 and 1920–90, Russia/USSR), Uzbekistan (1900–90, Russia/USSR), and East Timor (1976–99, Indonesia).

and of the world's population. The largest samples include 192 countries and over 6,500 observations. The smallest "full" samples include 149 countries and over 4,200 observations (see Tables 1 and 2). Even when all interpolated data for the dependent variable are dropped, the sample still includes 159 countries (in the minimal model) and over 2,600 observations (see Table 3, columns 1 and 2). Multiple imputation results include 199 countries and more than 6,200 observations (Table 2, columns 1–3). We include all countries for which data are available, including both developed and developing nations, in order to maximize variation on variables of theoretical interest and in order to capture both the successes and failures in human development around the world. (But see Table 3 for an analysis excluding the OECD and other groups of countries.) Most of our analyses incorporate Newey-West standard errors, with a one-period (AR1) correction for autocorrelation (exceptions are identified in the tables).

In all analyses, we lag the independent variables one time period. This separates the dependent variable from the predictors, offering some protection against X:Y endogeneity. Tests with longer (10 year) lags and tests that instrument for current values of all variables with previous values (in an Arellano-Bond format) provide further reassurance that endogeneity issues do not greatly affect the results.

Most analyses incorporate annual data (the unit of analysis is a country-year). Since data are available only at three-year intervals for some countries we interpolate missing data, as discussed above. In separate analyses, we conduct tests (a) with data observed at three-year intervals, and (b) with samples that exclude all interpolated data (see Tables 2 and 3).

Most analyses include country fixed effects. This technique removes many of the specification problems that typically plague cross-national studies, imposing a unique intercept for each country. This makes it less likely that the results suffer from omitted variable bias. The only exception would be a situation in which an unmeasured factor drives both (1) the change in the independent variable and (2) the change in IMR. In separate tests, we employ a series of spatial controls as a substitute for country fixed effects (see Table 2). However, because of the extreme uncertainty in model specification we have greater confidence in the fixed-effect format.

While fixed-effect regressions mitigate specification problems, they do not obviate them. In our search for dynamic (nonstatic) controls we try to identify factors that are measurable across national countries, are important influences on IMR, and are—at least to some extent—exogenous relative to the dependent variable (IMR) and our theoretical variable of interest (democracy). This search culminated in the choice of four control variables: *GDP per capita* (logged [World Bank 2003]), *urbanization* (World Bank 2003), *female illiteracy* (logit, World Bank 2003),[5] and *instability* (including assassinations, general strikes, guerilla warfare, government crises, purges, riots, revolutions, and antigovernment demonstrations).[6] We anticipate that IMR will show a negative relationship to GDP per capita and urbanization (Pritchett and Summers 1996) and a positive relationship to female illiteracy and instability. Of course, regime type may also have a causal impact on each of these controls. However, we assume that these effects are relatively minor and, more importantly, that any such effects should bias the direction of the regression results against our hypotheses. For example, if democracy influences growth it will have a salutary effect (Gerring et al. 2005); therefore, including per capita GDP in the regression models should underestimate the true effects of democracy on IMR. We also include a time-trend variable that begins in 1960 and increases by one in each subsequent year. Since trends in infant mortality rates are consistently sloped downwards—likely due to technological and related changes over time—this variable controls for the possibly spurious correlation between that variable and any similarly trended independent variables.

Once we remove country fixed effects the search for relevant controls becomes more onerous. In addition to the dynamic variables noted above, we search for static variables with the following characteristics: (1) they have been employed in previous studies; (2) they have a significant (presumably causal) impact on IMR in the benchmark model; (3) they are (at least partly) exogenous in the benchmark model; and

[5]The World Bank (2003) does not report illiteracy statistics for certain countries that have near 100% literacy rates. For these cases, we hand code a 0.5% female illiteracy rate. Illiteracy data for other countries are also somewhat limited. We impute missing data using the following technique. First, we fill in missing years between observed years of total adult illiteracy rates using linear interpolation. Second, we use annual regional averages to fill in missing data in the new interpolated total adult illiteracy variable. Third, we linearly interpolate missing years between observed years for female adult illiteracy. Fourth, we use the interpolated data for total illiteracy to impute values for interpolated female adult illiteracy. Fifth, because the data are bounded between 0 and 100, we take the logit of the imputed female illiteracy variable [logit=ln(x/(1-x))].

We also used multiple imputation to replace missing data in the entire dataset (see discussion in text), a process that employed the original World Bank variable.

[6]We standardize and add these variables (drawn from Banks 1994) together to form a composite index.

TABLE 1 Fixed-Effect Specification Tests

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
												Demo at t-10		0/1 coding
Democracy level	**-0.0004** (0.001)		**-0.0026**** (0.001)		**-0.0001** (0.001)		**0.0002** (0.001)	**-0.0012** (0.001)		**-0.0006***** (0.0002)		**-0.0085***** (0.001)		**-0.0260** (0.0159)
Democracy stock		**-0.0013***** (0.0001)		**-0.0012***** (0.0001)		**-0.0013***** (0.0001)	**-0.0013***** (0.0001)		**-0.0013***** (0.0001)		**-0.0001*** (0.00004)		**-0.0010***** (0.0001)	**-0.0125***** (0.001)
GDPpc (ln)			-0.3322*** (0.020)	-0.2591*** (0.020)	-0.2988*** (0.022)	-0.2373*** (0.021)	-0.2370*** (0.021)	-0.2800*** (0.022)	-0.2210*** (0.021)	-0.0118*** (0.004)	-0.0089** (0.004)	-0.3202*** (0.023)	-0.2539*** (0.023)	-0.2625*** (0.023)
Urbanization			-0.0031** (0.001)	-0.0075*** (0.001)	-0.0019 (0.001)	-0.0057*** (0.001)	-0.0057*** (0.001)	-0.0019 (0.001)	-0.0058*** (0.001)	-0.0006*** (0.0002)	-0.0010*** (0.0003)	-0.0020 (0.001)	-0.0061*** (0.001)	-0.0078*** (0.002)
Female illiteracy (logit)					-0.0058 (0.006)	-0.0025 (0.005)	-0.0024 (0.005)	0.0322*** (0.007)	0.0249*** (0.007)	-0.0051*** (0.001)	-0.0047*** (0.001)	-0.000006 (0.005)	0.0047 (0.005)	0.0133*** (0.005)
Instability					0.0014* (0.001)	0.0008 (0.001)	0.0008 (0.001)	0.0012 (0.001)	0.0006 (0.001)	0.000002 (0.0001)	-0.00002 (0.0001)	0.0009 (0.001)	0.0011 (0.001)	0.0010 (0.001)
Lagged dep var										0.9771*** (0.006)	0.9731*** (0.007)			
Trend	-0.0304*** (0.0005)	-0.0310*** (0.0004)	-0.0251*** (0.001)	-0.0238*** (0.001)	-0.0294*** (0.001)	-0.0268*** (0.001)	-0.0268*** (0.001)					-0.0293*** (0.001)	-0.0281*** (0.001)	-0.0237*** (0.001)
Year dummies								YES	YES					
Constant	4.4689*** (0.012)	4.4310*** (0.011)	6.9632*** (0.152)	6.5808*** (0.151)	6.7488*** (0.173)	6.4162*** (0.166)	6.4150*** (0.167)	6.6954*** (0.175)	6.3436*** (0.169)	0.1658*** (0.046)	0.1764*** (0.052)	6.9087*** (0.182)	6.5917*** (0.178)	607256*** (0.195)
Observations	*6569*	*6562*	*5547*	*5545*	*4495*	*4492*	*4492*	*4495*	*4492*	*4492*	*4489*	*4296*	*4291*	*3456*
Countries	*192*	*192*	*178*	*178*	*158*	*158*	*158*	*158*	*158*	*158*	*158*	*157*	*157*	*151*
Sample Period	*1960-00*	*1960-00*	*1960-00*	*1960-00*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-1994*
R-*square (within)*	*0.707*	*0.737*	*0.782*	*0.805*	*0.813*	*0.834*	*0.834*	*0.820*	*0.840*	*0.980*	*0.980*	*0.827*	*0.838*	*0.853*
Prob>F	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*

Dependent variable: IMR (ln). Unit of analysis: country-year. All independent variables are lagged one time period. Newey-West standard errors in parentheses.
Country fixed effects are included in each model. Models 5-7 are regarded as benchmark models.
*significant at 0.10 level; **significant at 0.05 level; ***significant at 0.01 level

TABLE 2 Further Specification Tests

	1 Annual	2 Annual	3 Annual	4 3-years	5 3-years	6 Annual	7 Annual	8 Annual	9 Annual	10 Annual	11 Annual
Democracy level	**-0.0021***		**-0.0004**	**-0.0005**		**-0.0095*****		**-0.0098*****		**0.0006**	
	(0.001)		(0.001)	(0.002)		(0.001)		(0.001)		(0.0005)	
Democracy stock		**-0.0010*****	**-0.0010*****		**-0.0012*****		**-0.0006*****		**-0.0006*****		**-0.0004*****
		(0.0001)	(0.0001)		(0.0002)		(0.00004)		(0.00004)		(0.00005)
GDPpc (ln)	-0.1721***	-0.1372***	-0.1376***	-0.3072***	-0.2474***	-0.3823***	-0.3496***	-0.3611***	-0.3357***	-0.0552***	-0.0451***
	(0.020)	(0.021)	(0.021)	(0.036)	(0.035)	(0.012)	(0.012)	(0.013)	(0.013)	(0.009)	(0.009)
Urbanization	-0.0013	-0.0035***	-0.0035***	-0.0018	-0.0055**	-0.0043***	-0.0043***	-0.0044***	-0.0043***	0.0009*	-0.0005
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female illiteracy (logit)	0.0127**	0.0118**	0.0115**	-0.0152	-0.0094	0.0486***	0.0392***	0.0826***	0.0660***	-0.0010	-0.0025
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.006)	(0.007)	(0.008)	(0.008)	(0.002)	(0.002)
Instability	0.0022**	0.0017**	0.0017**	0.0031*	0.0018	0.0087***	0.0077***	0.0087***	0.0078***	0.0003	0.0003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0003)	(0.0003)
Population (ln, 1960)						-0.0082	-0.0009	-0.0091*	-0.0021		
						(0.005)	(0.005)	(0.005)	(0.005)		
Ethnic fractionalize.						0.2771***	0.2800***	0.2712***	0.2757***		
						(0.034)	(0.033)	(0.034)	(0.033)		
Latitude (ln)						0.0336***	0.0382***	0.0344***	0.0379***		
						(0.012)	(0.011)	(0.012)	(0.011)		
Muslim						0.0026***	0.0022***	0.0021***	0.0018***		
						(0.0003)	(0.0003)	(0.0003)	(0.0003)		
Africa						0.2306***	0.2866***	0.2044***	0.2653***		
						(0.033)	(0.032)	(0.032)	(0.032)		
Asia						-0.1667***	-0.1609***	-0.1829***	-0.1737***		
						(0.032)	(0.033)	(0.032)	(0.032)		
Latin America						0.3516***	0.2890***	0.3416***	0.2850***		
						(0.027)	(0.026)	(0.026)	(0.025)		
Socialism						-0.1869***	-0.3632***	-0.1412***	-0.3146***		
						(0.038)	(0.043)	(0.038)	(0.044)		
Lagged dep var										0.8644***	0.8186***
										(0.009)	(0.010)
Trend	-0.0278***	-0.0278***	-0.0278***	-0.0285***	-0.0259***	-0.0239***	-0.0255***				
	(0.001)	(0.001)	(0.0009)	(0.001)	(0.001)	(0.001)	(0.001)				
Multiple imp.	YES	YES	YES								

TABLE 2 (Continued)

	1	2	3	4	5	6	7	8	9	10	11
	Annual	**Annual**	**Annual**	**3-years**	**3-years**	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**	**Annual**
Year dummies								YES	YES		
Country fe	YES	YES	YES	YES	YES					YES	YES
Constant	5.7715***	5.5923***	5.5939***	6.7762***	6.4515***	7.3107***	7.0463***	7.3441***	7.0117***	-0.0043***	-0.0048***
	(0.159)	(0.169)	(0.169)	(0.282)	(0.274)	(0.107)	(0.104)	(0.112)	(0.108)	(0.0004)	(0.0004)
Observations	*6233*	*6233*	*6233*	*1640*	*1639*	*4371*	*4368*	*4371*	*4368*	*4328*	*4325*
Countries	*199*	*199*	*199*	*158*	*158*	*149*	*149*	*149*	*149*	*157*	*157*
Sample Period	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*
R-square (within)	*0.732*	*0.749*	*0.749*	*0.813*	*0.833*	*0.882*	*0.891*	*0.886*	*0.893*		
Prob>F	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*		
Sargan test (prob)										*< 0.0001*	*< 0.0001*

Dependent variable: IMR (ln). All independent variables are lagged one time period. Newey-West standard errors in parentheses and AR1 correction for serial autocorrelation, except in the final two models, where the analysis is Arellano-Bond.
* significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level

(4) data are widely available over the sample period. These criteria help to identify several additional control variables used in the nonfixed effects estimations: *population* in 1960 (ln, World Bank 2003),[7] *ethnic fractionalization* (the likelihood that two persons randomly chosen from a population belong to different ethnic groups; Alesina et al. 2003), *latitude* (absolute value of distance from the equator, logged, LaPorta et al. 1999), *Muslim* (percent Muslims; CIA *World Factbook* [online]), *Africa* (dummy), *Asia* (dummy), *Latin America* (dummy), and *Socialist legal origin* (La Porta et al. 1999). We anticipate positive signs for population, ethnic fractionalization, Muslim, Africa, and Latin America, and negative signs for latitude, Asia and Socialism. We have no doubt that other factors could be proposed, and some of those chosen here might be questioned, either on theoretical or empirical grounds. However, we doubt that a different selection of controls would alter the substance of the results reported here. In any event, we regard these tests as robustness checks for the fixed-effect models. The online appendix provides descriptive statistics for all variables.

Results

To reiterate, our twin research questions concern the possible causal effects of contemporary democracy (a level variable) and democratic history (a stock variable) on variation in infant mortality rates. We hypothesize that a country's stock of democracy, but not its current regime status, will be associated with a lower rate of infant mortality in the following period, all other things being equal. In the following tables, we present these two sets of results side by side.

Tables 1, 2, and 3 present our findings. In each case, the fit of the models is quite good, with F-tests significant at the 0.0001 level and high R^2 values.[8] Table 1 displays a series of fixed-effect regressions with different model specifications. Models 1 and 2 include a minimal model including only the democracy level and stock variables, respectively, and a time-trend control. Models 3 and 4 introduce two additional control variables (GDP per capita and urbanization) to capture the effect of economic development. Models 5, 6, and 7

[7]We treat population as a static variable, measured in the first year of the analysis, in order to minimize endogeneity problems.

[8]This R^2 is harvested from the first "phase" of Newey-West regressions, before the error correction process. Note that the use of fixed effects and the various time-control variables inflates the R^2 values obtained here. We report them as a measure of fit for the interested reader, without placing much substantive emphasis on them.

TABLE 3 Sample Restrictions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Excluding	Interpolated data		Most autocratic		Most democratic		Asia		Latin America		Africa		Middle East		OECD	
Democracy level	**-0.0015**		**-0.0005**		**-0.0007**		**-0.0028****		**-0.0028***		**-0.0010**		**-0.0032****		**0.0040*****	
	(0.001)		(0.001)		(0.001)		(0.001)		(0.002)		(0.002)		(0.001)		(0.001)	
Democracy stock		**-0.0013*****		**-0.0012*****		**-0.0011*****		**-0.0013*****		**-0.0013*****		**-0.0005*****		**-0.0016*****		**-0.0008*****
		(0.0001)		(0.0001)		(0.0001)		(0.0001)		(0.0001)		(0.0001)		(0.0001)		(0.0001)
GDPpc (ln)	-0.3343***	-0.2686***	-0.3225***	-0.3393***	-0.2674***	-0.2705***	-0.3335***	-0.2303***	-0.3552***	-0.2711***	-0.2105***	-0.2011***	-0.3826***	-0.2994***	-0.2098***	-0.1989***
	(0.027)	(0.025)	(0.027)	(0.030)	(0.028)	(0.028)	(0.027)	(0.025)	(0.031)	(0.030)	(0.032)	(0.031)	(0.028)	(0.027)	(0.026)	(0.026)
Urbanization	-0.0011	-0.0053***	0.0024	-0.0031*	-0.0015	-0.0063***	0.0009	-0.0038**	-0.0006	-0.0059***	-0.0082***	-0.0095***	0.0024	-0.0030*	-0.0004	-0.0016
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female illiteracy	-0.0098	-0.0094	-0.0124	-0.0054	-0.0095	-0.0115*	-0.0405***	-0.0307***	-0.0193*	-0.0165	0.0285***	0.0253***	-0.0286***	-0.0327***	-0.0104	-0.0093
	(0.008)	(0.007)	(0.009)	(0.008)	(0.007)	(0.007)	(0.009)	(0.009)	(0.011)	(0.011)	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)
Instability	0.0017*	0.0009	0.0022**	0.0006	-0.0006	0.0015	0.0021**	0.0014	0.0026**	0.0013	0.0013	0.0011	0.0019*	0.0007	-0.0002	-0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Trend	-0.0317***	-0.0283***	-0.0335***	-0.0281***	-0.0278***	-0.0271***	-0.0331***	-0.0296***	-0.0311***	-0.0275***	-0.0350***	-0.0332***	-0.0315***	-0.0270***	-0.0295***	-0.0281***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	6.8436***	6.5566***	6.5524***	7.0506***	6.5248***	6.6090***	6.7179***	6.1910***	6.9191***	6.5549***	6.2356***	6.2398***	6.9744***	6.6223***	6.1426***	6.0239***
	(0.225)	(0.209)	(0.235)	(0.244)	(0.207)	(0.205)	(0.239)	(0.214)	(0.258)	(0.237)	(0.274)	(0.263)	(0.246)	(0.222)	(0.190)	(0.190)
Observations	*2603*	*2602*	*2426*	*2400*	*1715*	*1832*	*2290*	*2289*	*2247*	*2246*	*2084*	*2084*	*2390*	*2389*	*1607*	*1606*
Countries	*159*	*159*	*153*	*146*	*134*	*141*	*139*	*139*	*138*	*138*	*113*	*113*	*142*	*142*	*128*	*128*
Sample Period	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*	*1960-99*
R-square (within)	*0.850*	*0.866*	*0.856*	*0.873*	*0.771*	*0.801*	*0.855*	*0.871*	*0.849*	*0.866*	*0.901*	*0.903*	*0.859*	*0.881*	*0.753*	*0.761*
Prob>F	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*	*< 0.0001*

Dependent variable: IMR (ln). Unit of analysis: country-year (with fixed effects). All independent variables are lagged one time period. Newey-West standard errors in parentheses. AR1 correction for serial autocorrelation. Exclusions from the full sample are explained in the text.
*significant at 0.10 level; ** significant at 0.05 level; *** significant at 0.01 level

add two additional controls: female illiteracy and a variable intended to measure political conditions in a country that might have strong effects on human development: the level of political instability. We regard columns 5, 6, and 7 as benchmark models, since they include a full selection of plausible controls in a fixed-effects format with a time trend and correction for autocorrelation. Model 7 introduces democracy level and stock together in the same estimation.

Models 8 and 9 substitute *T-1* annual year dummies for the trend variable to provide another means of modeling change over time. Models 10 and 11 provide yet another approach to modeling time effects, this time with a lagged dependent variable (and fixed effects, resulting in a Least Squares Dummy Variable, or LSDV, estimator; see Beck and Katz 2004). The coefficient for the lagged dependent variable is quite high (0.97–0.98), which is not surprising given the strongly trended nature of IMR data. The inclusion of the lagged dependent variable means that the coefficients for the remaining variables capture only their short-term effects. To calculate their long-term effects, we divide the coefficient by one minus the coefficient of the lagged dependent variable (Beck and Katz 2004). This generates a long-run coefficient of -0.028 for the level variable and -0.003 for the stock variable. Models 12 and 13 revert to the benchmark equation (employed in models 5 and 6), but lag the two democracy variables by 10 years (instead of the usual single-year lag). This addresses the possibility that the causal relationship between democracy and human development may have a long time lag.

Model 14 replicates the benchmark model of column 7 with a different coding of democracy. Rather than our usual continuous measure (drawn from the Polity database), we adopt a binary coding of this variable drawn from Boix and Rosato (2001). A stock measure is constructed using the same 1% depreciation rate employed for our usual variable. Results are quite similar to our baseline results showing that the findings are robust to different methods of measuring democracy.

Table 2 provides further specification tests. Models 1–3 employ multiple imputation to address concerns over the possible biases introduced by missing data (King et al. 2001), a procedure discussed above. Smaller, poorer countries tend to have less data available, and the exclusion of those nonrandomly missing cases could result in a biased sample (Ross 2006). Reassuringly, results for the fully imputed dataset are quite similar to those for our narrower sample. Models 4 and 5 test the benchmark equation on data drawn from three-year (rather than annual) intervals. This accords with the infrequent nature of data collection in many developing countries. Reassuringly, results are quite similar to those generated from our annual samples. Models 6 and 7 revert to annual data but include a series of static controls in place of country fixed effects. Models 8 and 9 replicate 6 and 7 except that they include annual dummy variables instead of the trend variable. Finally, models 10 and 11 test the benchmark equation using the Arellano-Bond technique (Arellano and Bond 1991). This method combines first differencing with a series of lags, equivalent to the total number of prior observations in the dataset, for each variable in the model.

Table 3 imposes a series of restrictions on the full sample. Models 1 and 2 exclude all data that were interpolated for the dependent variable (as described in a previous section); it thus represents the data as drawn from the WDI dataset, without further additions. Models 3 and 4 exclude the decile of most autocratic countries in the full sample, while models 5 and 6 exclude the most democratic decile, thus eliminating the potential impact of extreme cases on either end of the democracy continuum.[9] Subsequent models exclude various regions and groups of countries around the world: Asia (models 7 and 8), Latin America (9 and 10), Africa (11 and 12), Middle East (13 and 14), and the OECD (15 and 16).

With a few exceptions, most of the control variables perform as expected. Higher levels of economic development, lower rates of female illiteracy and less political instability are generally associated with lower rates of infant mortality. For the key theoretical variables (in bold), the patterns displayed in Tables 1–3 are striking.

Arguably, the most important test is contained in Table 1, model 7, which includes the level and stock variables together in the same benchmark model. We find that the coefficient and standard error for the stock variable is unaffected, while the level variable shows a *positive* (though not significant) relationship to IMR. (These results are stable when we lag the stock variable by two years so as to better disentangle its effects from the democracy-level variable, which has a one-year lag.) This inclines us to conclude that long-term democratic stock, not the status of the current regime, causes human development. In other tests the democracy-level variable also performs inconsistently. In some specifications it is associated

[9]Tests (not reported) that exclude both the most *and* least democratic cases together do not differ appreciably from the results presented here.

with lower numbers of infant deaths but the relationship reaches conventional levels of statistical significance in the expected direction only occasionally.

The inconsistency of the findings for democracy level suggests that the true relationship between democracy and IMR is a historically mediated one. By contrast, results for the democracy stock variable are robust in each and every specification (over 40 in total reported in Tables 1–3), at the 0.01 level of significance or better in all cases but one, where it attained the 0.10 level ($p=0.059$).

Importantly, this way of measuring democracy is less subject to certain identification problems. Both democracy (at *t-1*) and IMR (at *t*) may be caused by some underlying factor that is unmeasured. If so, these models are misidentified. It is less likely that a stock measure of democracy will be subject to this sort of problem since year-to-year changes in this variable are a product of its (very long) history. An additional year of high-quality democracy matters more for Benin (a new democracy) than for the United States (an old democracy). This means that the slope of the democracy stock variable is quite different from the slope of the democracy level variable for any given country. Granted, a country's accumulated stock of democracy is also the product of underlying causes; it is not an unmoved mover. Yet, insofar as these underlying causes comprise static features of a country—e.g., its prior history, demography, geography, or culture—they are captured in the country fixed effect and do not affect those results. Our purpose here is to test the effect of democratic stock on human development. The prior question of what causes democratic stock lies beyond the scope of this inquiry.

Beyond the specification tests presented in Tables 1–3, an additional battery of tests is presented in the online appendix (Table A5). First, it may be that our democracy stock variable may be capturing regime durability rather than democracy per se. To control for this, we include a variable that measures regime volatility (a 10-year moving variance of Polity2). Concerned that perhaps democracy might be proxying some other good governance factor, we tested four additional models that add various governance indicators into our benchmark model: (1) rule of law (World Bank), (2) corruption (World Bank), (3) regulatory quality (World Bank), and (4) corruption (Political Risk Services). Finally, we tested a model that codes contiguous colonies pre-independence Polity2 scores as zero rather than coding them as the same as their colonizer. In each of these six robustness checks, our results were not disturbed.

We also conducted a series of additional tests for potential confounders. First, we used a dependent variable that measures countries' deviations from the global infant mortality mean in a particular year (King and Zeng 2001). This is a detrending exercise, intended to focus attention on "best" and "worst" performers relative to the global trend.

Second, we included a control for a country's real rate of per capita economic growth at time *t* (as opposed to *t-1*, the lag used with other independent and control variables). This should control for the state of the economy in a given year, and, to the extent that economic patterns within a given country reflect broader tendencies, broader global trends that we might not otherwise capture.

Third, we employed an interaction term between democratic stock and GDP per capita, to further test for the possibility that democratic history matters differently for countries at varying levels of development. (Plausibly, stock matters more for poorer countries than for richer ones.) Results for the interaction term are not significant, nor does its inclusion affect the performance of the stock variable.

Fourth, as noted previously, there are a variety of ways in which one might calculate, and test, the concept of democratic stock. The 1% depreciation rate employed in this study is intended to capture the possibility that democratic stock does not accumulate linearly over time; it is not intended, however, to serve as a definitive measurement of that concept. Regrettably, because of the highly trended nature of democracy (countries that are highly democratic today are likely to have maintained this pattern in previous years) it is not possible to employ a distributed lag model; the multiple lags would be so highly correlated as to introduce prohibitive problems of collinearity. However, we did experiment with alternative measurements of democratic stock, including both a quadratic and square root measure of this key variable. In both cases, results are substantively the same as those presented here. We also tested a 5% depreciation measure of democratic stock, with the same substantive results. We postpone further exploration of the stock concept for future work. In each of these additional tests (not reported), our central results hold.

Finally, it is important to note that the relationship between democratic stock and infant mortality observed here is not only statistically significant but also substantively important. Consider the following projection, based on the coefficients reported for democratic stock in models 6 and 7 in Table 1 (-0.0013). For a country with no prior accumulated democratic stock (e.g., a new country), 10 years of

full democracy (Polity2=10) translates into a 10% reduction in IMR. By way of comparison, a similar improvement in human development, if driven solely by economic processes, would require a 40% increase in GDP per capita over the course of a decade—an extraordinarily rare accomplishment, except in the case of the sudden discovery of mineral wealth (itself often associated with pathological effects). Ten-point improvements in the Polity (21-point) index are a good deal more common than 40% increases in per capita GDP. Thus, the results obtained in this study suggest that regime type has a substantial effect on human development. The wave of democratization felt across many regions of the developing world over the past several decades may well be an important ingredient in contemporary achievements in infant mortality over the same period.

Discussion

Contrary to much recent work, this article argues that there is no strong or robust relationship between a country's current regime type and its subsequent human development, as measured by infant mortality rates. In this respect, we agree with recent critiques of the received view (Gauri and Khaleghian 2002; McGuire 2004; Ross 2006; Shandra et al. 2004). However, we argue that a robust causal relationship does appear if democracy is considered as a long-run, historical phenomenon. Regression tests indicate that a stock measure of democracy is associated with improved human development. Moreover, the construction of this variable may also make it less vulnerable to some of the potential identification problems that affect regression tests using a "level" measure of democracy.

This is a new angle on an old question. Indeed, we find only one published article— Ross (2006)—that purports to test the relationship between democratic history and infant mortality. Ross's much-cited article focuses mostly on the relationship between democracy level and mortality; here, our conclusions coincide. Ross also provides several cross-national tests using a stock indicator of democracy, with null findings that contradict our own. Here, some discussion is in order (for further details see online appendix, Table A6).

We were able to replicate Ross's reported results in most respects.[10] What bears careful review are the choices Ross makes in estimation strategy. First, Ross includes a lagged dependent variable in most estimations, introducing the possibility of unit root problems and of confounding by way of colliders (Clarke 2005). Second, he observes data at five-year periods from 1970 to 2000, providing a six-period panel in most instances (one period is dropped when a lagged dependent variable is included). This is a fairly short panel, especially given the trended nature of the key variables and heterogeneity across units (countries). Arguably, annual data observed over a somewhat longer period (our panel begins in 1960) provide a more sensitive read of the time-varying relationship between regime type and human development. Third, Ross chose a binary measure of democracy to compute a country's democratic history (stock = years democratic, logged). This approach to measurement presumes that there is no difference in causal impact between hard autocracies (e.g., North Korea) and soft authoritarian states (e.g., Mexico under the Institutional Revolutionary Party) or between semi-democratic polities (e.g., Malaysia since 1969) and strongly democratic polities (e.g., Mauritius). All cases are lumped into one of two piles. While this sort of dichotomous coding may be theoretically defensible for certain purposes, it is inconsistent with theoretical arguments about democracy and outcomes like infant mortality. Here, the presumed intermediary factors (in Ross's account and our own) are matters-of-degree, rather than either/or. Dichotomous coding also introduces potentially large measurement errors, for misclassifications have extreme effects where there are only two possible coding categories.

To be sure, Ross's choices in estimation and measurement are not entirely implausible. Indeed, we employ most of them in various robustness tests (see Tables 1–3 and A6). However, we regard them as less plausible than the alternatives, for reasons stated briefly above and in the text. They are also highly consequential.

When one substitutes a trend variable for the lagged dependent variable, one finds highly significant results for the democratic stock variable. When one substitutes our stock variable (computed from a continuous rather than dichotomous measure of democracy) into Ross's estimations with fixed effects and period dummies (no lagged dependent variable), one finds similarly significant results. Only when using Ross's less-sensitive measure of democratic stock *or* a combination of a lagged dependent variable and data observed at five-year intervals does the stock variable fail to achieve statistical significance.

[10]Ross only indicates levels of significance at the 5% level or better. Two of our replications of his estimations find results statistically significant at the 10% level.

Numerous methods of modeling temporal autocorrelation in our annual data—with a time trend, lagged dependent variable, period dummies, first differences, or dynamic models (Arellano Bond)—yield highly significant results (see Tables 1 and A6). Thus, only in one very specific combination of modeling choices are we able to replicate Ross's null findings using our measure of democratic stock. Since we do not view this particular model as the most likely representation of the data-generating process we see the story elucidated by the cross-national data quite differently from Ross.

By way of conclusion, it may be observed that the relationship between democracy stock and lower infant mortality is much more robust than many other cross-national data patterns commonly regarded as causal, e.g., human capital and growth (Barro 1991) or trade and growth (Rodriguez and Rodrik 2001). While nothing is certain in the business of observational data analysis, the associational patterns between stock and IMR illustrated in Tables 1–3 and in additional tables in our online appendix offer strong circumstantial grounds for causal inference.

Moreover, there are strong *theoretical* grounds for expecting a causal relationship to exist between democratic stock and human development, as sketched in the first part of the article. Looking back at the proposed causal pathways—competition-induced accountability, a vigorous civil society, a culture of equality, and institutionalization—one can appreciate that such factors are unlikely to take effect immediately. Indeed, it may be unreasonable to expect any relationship between regime type and distal policy outcomes such as infant mortality to materialize in the space of a year, or even in five- or ten-year periods, as stipulated (often only implicitly) by extant studies. It is a country's regime history, more than its present status that determines whether, and to what extent, that society is likely to achieve significant improvements in the lives and livelihoods of its citizens. This, we argue, is likely because long-term democracies benefit from more political competition leading to greater accountability, stronger civil societies pushing for and promoting human development, the development of norms that support greater demands for equality, and higher levels of institutionalization, relative to authoritarian regimes or new democracies.

We have not attempted to test the relative importance of each of the posited causal pathways. This represents a difficult conceptual and empirical challenge, as plausible causal mechanisms are diverse, overlapping, and in most instances resistant to operationalization (Gerring 2011). Nonetheless, we hope that future work on this topic will shed light on the explanatory factors that may be at work within the distal relationship that we have explored here.

If the logic of our argument is correct, it may also apply to other measures of development, a matter that we are currently exploring in a companion study (Gerring and Thacker [in process]). Arguably, most, and perhaps all political-institutional variables are time-dependent, which is to say that their effects today are a product, in part, of their histories. These histories may be quite long, and quite consequential. Figuring out the ways in which present outcomes depend upon past choices is a critical task for every empirical research technique, not simply a matter to be reserved for case-study or historical-institutionalist researchers, though we surely have much to learn from work in these historically oriented genres (Collier and Collier 1991; Mahoney and Rueschemeyer 2003; Pierson 2004).

The practical implications of this argument introduce grounds for both optimism and caution with respect to the ability of developing countries to improve their levels of human development. Realistically, countries should not expect large immediate dividends in human development to result from democratic transitions. On the other hand, given sufficient time, democracy should begin to yield important, tangible benefits to the underprivileged in society. In a world characterized by chronically short time horizons, the substantial political challenge is to allow democratic institutions the time necessary to realize these persistent but distal benefits.

Acknowledgments

We are grateful for support for this research from the Clinton Global Initiative and the Frederick S. Pardee Center for the Study of the Longer-Range Future, both at Boston University. We thank Julián Arévalo for his excellent research assistance and Michael Ross for generously providing his data and comments on an earlier version. The usual disclaimers apply.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8 (2): 155–94.

Alvarez, Sonia, Evelina Dagnino, and Arturo Escobar, eds. 1998. *Cultures of Politics/Politics of Cultures: Revisioning Latin American Social Movements*. Boulder, CO: Westview.

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an

Application to Employment Equations.'' *Review of Economic Studies* 58 (April): 277–97.

Banks, Arthur S. 1994. ''Cross-National Time-Series Data Archive.'' Center for Social Analysis, State University of New York at Binghamton.

Barro, Robert J. 1991. ''Economic Growth in a Cross-section of Countries.'' *Quarterly Journal of Economics* 106 (May): 407–33.

Beck, Nathaniel, and Jonathan N. Katz. 2004. ''Time-Series—Cross-Section Issues: Dynamics, 2004.'' New York University and California Institute of Technology. Unpublished manuscript.

Boix, Carles. 2001. ''Democracy, Development and the Public Sector.'' *American Journal of Political Science* 45 (January): 1–17.

Boix, Carles, and Sebastian Rosato. 2001. ''A Complete Data Set of Political Regimes, 1800–1999.'' Chicago: University of Chicago.

Bollen, Kenneth A. 1993. ''Liberal Democracy: Validity and Method Factors in Cross-National Measures.'' *American Journal of Political Science* 37 (November): 1207–30.

Bollen, Kenneth A., and Pamela Paxton. 2000. ''Subjective Measures of Liberal Democracy.'' *Comparative Political Studies* 33 (February): 58–86.

Bowman, Kirk, Fabrice Lehoucq, and James Mahoney. 2005. ''Measuring Political Democracy: Case Expertise, Data Adequacy, and Central America.'' *Comparative Political Studies* 38 (October): 939–70.

Brown, David S., and Ahmed Mushfiq Mobarak. 2009. ''The Transforming Power of Democracy: Regime Type and the Distribution of Electricity.'' *American Political Science Review* 103 (May): 193–214.

Brown, David S., and Wendy Hunter. 2004. ''Democracy and Human Capital Formation: Education Spending in Latin America, 1980–1997.'' *Comparative Political Studies* 37 (September): 842–64.

Bueno de Mequita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Clarke, Kevin A. 2005. ''The Phantom Menace: Omitted Variable Bias in Econometric Research.'' *Conflict Management and Peace Science* 22 (4): 341–52.

Collier, Ruth Berins, and David Collier. 1991. *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Princeton, NJ: Princeton University Press.

Dreze, Jean, and Amartya Sen. 1989. *Hunger and Public Action*. Oxford: Clarendon Press.

Filmer, Deon, and Lant Pritchett. 1999. ''The Impact of Public Spending on Health: Does Money Matter?'' *Social Science and Medicine* 49 (November): 1309–23.

Gauri, Varun, and Evan Lieberman. 2006. ''Boundary Politics and Government Responses to HIV/AIDS in Brazil and South Africa.'' *Studies in Comparative International Development* 41 (3): 47–73.

Gauri, Varun, and Peyvand Khaleghian. 2002. ''Immunization in Developing Countries: Its Political and Organizational Determinants.'' *World Development* 30 (December): 2109–32.

Gerring, John. 2011. ''Causal Mechanisms: Yes, But . . . '' *Comparative Political Studies* 44 (11): 1499–1526.

Gerring, John, Philip Bond, William Barndt, and Carola Moreno. 2005. ''Democracy and Growth: A Historical Perspective.'' *World Politics* 57 (April): 323–64.

Gerring, John, and Strom Thacker. In process. *Democracy and Development: A Historical Perspective*. Boston University.

Ghobarah, Hazem Adam, Paul Huth, and Bruce Russett. 2004. ''Comparative Public Health: The Political Economy of Human Misery and Well-Being.'' *International Studies Quarterly* 48 (1): 73–94.

Haggard, Stephan. 1991. ''Inflation and Stabilization.'' *Politics and Policy Making in Developing Countries: Perspectives on the New Political Economy*, ed. Gerald M. Meier. San Francisco: ICS Press, 233–49.

Hardin, Russell. 1999. *Liberalism, Constitutionalism, and Democracy*. Oxford: Oxford University Press.

Hill, Kenneth, Rohine Pande, Mary Mahy, and Gareth Jones. 1999. *Trends in Child Mortality in the Developing World: 1960–1996*. New York: UNICEF.

Hite, Katherine, and Paola Cesarini, eds. 2004. *Authoritarian Legacies and Democracy in Latin America and Southern Europe*. Notre Dame, IN: University of Notre Dame Press.

Huntington, Samuel P. 1968. *Political Order in Changing Societies*. New Haven, CT: Yale University Press.

Kane, Penny. 1989. ''Famine in China 1959–61: Demographic and Social Implications.'' In *Differential Mortality: Methodological Issues and Biosocial Factors*, ed. Lado Ruzicka, Guillaume Wunsch, and Penny Kane. Oxford: Clarendon Press, 231–53.

Kapstein, Ethan B., and Nathan Converse. 2008. *The Fate of Young Democracies*. Cambridge: Cambridge University Press.

Keefer, Philip. 2006. ''Clientelism, Credibility and the Policy Choices of Young Democracies.'' *American Journal of Political Science* 51 (October): 804–21.

Kudamatsu, Masayuki. Forthcoming. ''Has Democratization Reduced Infant Mortality in sub-Saharan Africa? Evidence from Micro Data.'' *Journal of the European Economic Association*.

King, Gary, James Honaker, and Anne Joseph, Kenneth Scheve. 2001. ''Analyzing Incomplete Political Science Data.'' *American Political Science Review* 95 (March): 49–69.

King, Gary, and Langche Zeng. 2001. ''Research Note: Improving Forecasts of State Failure.'' *World Politics* 53 (July): 623–58.

Kurzman, Charles K., Regina W. Werum, and Ross E. Burkhart. 2002. ''Democracy's Effect on Economic Growth: A Pooled Time-Series Analysis, 1951–1980.'' *Studies in Comparative International Development* 37 (Spring): 3–33.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1999. ''The Quality of Government.'' *Journal of Economics, Law and Organization* 15 (1): 222–79.

Lake, David A., and Matthew A. Baum. 2001. ''The Invisible Hand of Democracy: Political Control and the Provision of Public Services.'' *Comparative Political Studies* 34 (August): 587–621.

Lenski, Gerhard. 1966. *Power and Privilege: A Theory of Social Stratification*. Chapel Hill: University of North Carolina Press.

Levitsky, Steven. 1998. ''Institutionalization and Peronism: The Concept, the Case, and the Case for Unpacking the Concept.'' *Party Politics* 4 (1) 77–92.

Lipset, Seymour Martin. 1959. ''Some Social Requisites of Democracy: Economic Development and Political Development.'' *American Political Science Review* 53 (March): 69–105.

Mahoney, James, and Dietrich Rueschemeyer, eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Marshall, Monty G., and Keith Jaggers. 2000. ''Polity IV Project: Political Regime Characteristics and Transitions, 1800–2000.'' http://www.systemicpeace.org/polity/polity4.htm.

McGuire, James W. 2004. ''Democracy, Health Care, and Under-5 Mortality: A Cross-National Study of Developing Countries.'' Wesleyan University. Unpublished manuscript.

McGuire, James W. 2010. *Wealth, Health, and Democracy in East Asia and Latin America*. Cambridge: Cambridge University Press.

Meltzer, Allan H., and Scott F. Richard. 1981. ''A Rational Theory of the Size of Government.'' *Journal of Political Economy* 89 (5): 914–27.

Muller, Edward N. 1988. ''Democracy, Economic Development, and Income Inequality.'' *American Sociological Review* 53 (February): 50–68.

Munck, Gerardo L., and Jay Verkuilen. 2002. ''Measuring Democracy: Evaluating Alternative Indices.'' *Comparative Political Studies* 35 (1): 5–34.

Olson, Mancur. 1982. *The Rise and Decline of Nations*. New Haven, CT: Yale University Press.

Parker, Richard. 1994. ''Policy, Activism, and AIDS in Brazil.'' In *Global AIDS Policy*, ed. Douglas A. Feldman. Westport, CT: Bergin and Garvey, 28–46.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Piven, Frances Fox, and Richard A. Cloward. 1977. *Poor People's Movements: Why They Succeed, How they Fail*. New York: Pantheon.

Polsby, Nelson. 1968. ''The Institutionalization of the U.S. House of Representatives.'' *American Political Science Review* 62 (1): 144–68.

Pritchett, Lant, and Lawrence H. Summers. 1996. ''Wealthier is Healthier.'' *Journal of Human Resources* 31 (Fall): 841–68.

Przeworski, Adam, Michael Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.

Riskin, Carl. 1995. ''Food, Poverty, and Development Strategy in the People's Republic of China.'' In *Hunger in History: Food Storage, Poverty, and Deprivation*, ed. Lucile F. Newman et al. Cambridge: Blackwell, 331–52.

Rodriguez, Francisco, and Dani Rodrik. 2001. ''Trade Policy and Economic Growth: A Skeptic's Guide to Cross-National Evidence.'' In *NBER Macroeconomics Annual 2000*, ed. Ben Bernanke and Ken Rogoff. Cambridge, MA: MIT Press, 261–338.

Ross, Michael. 2006. ''Is Democracy Good for the Poor?'' *American Journal of Political Science* 50 (October): 860–74.

Rubin, Jeffrey W. 1997. *Decentering the Regime: Ethnicity, Radicalism, and Democracy in Juchitán, Mexico*. Durham, NC: Duke University Press.

Shandra, John M., Jenna Nobles, Bruce London, and John B. Williamson. 2004. ''Dependency, Democracy, and Infant Mortality: A Quantitative, Cross-National Analysis of Less Developed Countries.'' *Social Science and Medicine* 59 (July): 321–33.

Sondhi, Sunil. 2000. ''Combating Corruption in India: The Role of Civil Society.'' Prepared for the 18th World Congress of International Political Science Association (August 1–5), Quebec City, Canada.

Treier, Shawn, and Simon Jackman. 2008. ''Democracy as a Latent Variable.'' *American Journal of Political Science* 52 (January): 201–17.

Vanhanen, Tatu. 1990. *The Process of Democratization: A Comparative Study of 147 States, 1980–88*. New York: Crane Russak.

World Bank. 2003. *World Development Indicators CD-ROM*. Washington, DC: World Bank.

John Gerring is a Professor of Political Science at Boston University, Boston, MA 02215.

Strom C. Thacker is a Professor of International Relations and Political Science and Associate Dean of the Faculty, Social Sciences at Boston University, Boston, MA 02215.

Rodrigo A. Alfaro is Principal Economist at BBVA Research, Santiago, Chile.