The *Journal of Public and International Affairs* is a joint publication of the Association of Professional Schools of International Affairs (APSIA) and the Woodrow Wilson School of Public and International Affairs at Princeton University. It seeks to publish both scholarly and expository articles on a diverse range of subjects, covering the areas of international affairs, development studies, and domestic policy. The journal is an annual publication, accepting submissions from any advanced student in the field from any accredited institution. Submissions are reviewed in a blind screening process by an editorial board composed of students at Princeton and by contributing editors from the other APSIA schools. Those wishing to submit papers for consideration may send them to: *JPIA*, Robertson Hall, Woodrow Wilson School, Princeton University, Princeton, NJ 08544-1013. A subscription and comments form can be found on the last page of this issue.
CONTENTS

Editors' Note v

1 States or Markets? The Reform of Health Insurance in Chile, 1981–1994 1
Norbert R. Schady

Timothy William Waters

3 Child-Support Laws and Contraceptive Care 47
Timothy Irwin

4 One Man, One Grant: Equity in Federal Aid to States 60
Karen Baehler

5 Manipulation of Islam by the State: Islam as an Official Nationalism in Pakistan, Malaysia, and Uzbekistan 86
Laurence S. Hanauer

6 Central Bank Reform in Latin America: Will Independence Guarantee Low Inflation? 107
John F. Normand

7 Asian Values or Asian Ideology? The Role of Culture in the International Relations of East Asia 129
Andrew Oros

8 Fiscal Stress and Urban Regimes: Straining the Capacity to Govern 148
Preston Niblack

9 The Intelligence Authorization Act of 1991 and Its Implications for Congressional Oversight of Covert Action 171
Ladd Everitt

10 Most-Favored-Nation Status and China: An Economic Analysis of Linkage 188
Phillip Saunders
Editors' Note

The Journal of Public and International Affairs (JPIA) is entering its sixth year as the only national journal for graduate students working in public and international affairs, and as such serves as a unique forum for the views of future scholars and policymakers.

For the 1995 issue of the JPIA, contributing editors at thirteen APSIA schools in the United States and Canada helped gather approximately 110 submissions for the selection process. Thirteen of these editors traveled to Princeton for our annual Reading Weekend, making it our largest yet. The editorial board used a blind grading process to select the ten papers published herein.

The ten papers selected cover a range of diverse topics, from Norbert Schady's article on the health-insurance system in Chile, to Ladd Everitt's examination of congressional oversight of covert action. Two papers address issues of culture and ethnicity in the formation of national identity—Timothy Waters' discussion of a Gypsy national identity, and Larry Hanauer's case study of Islam as a force behind nationalist movements in Pakistan, Malaysia, and Uzbekistan. Tim Irwin's paper, on the other hand, takes a more theoretical approach, using economic techniques to analyze the implications of alternative child-support policies. Karen Baehler and Preston Niblack discuss problems with a large domestic impact—equity in providing federal aid to states, and the role of fiscal stress in city governance.

The JPIA staff is grateful for the continued support of the APSIA deans and contributing editors. In particular, we would like to thank Kay King, APSIA's executive director, for her gracious help and enthusiasm for the journal.

The editorial board would like to extend special thanks to the Woodrow Wilson School and Princeton University for their long-standing support of the JPIA. In particular, we wish to thank Acting Dean James Trussell and Assistant Dean Sybil Stokes, for their generous financial and planning assistance; Agnes Pearson, for helping with the logistics of the selection process; and Pat Coen, for her ongoing assistance in the production of the journal.
This paper analyzes the creation of private health insurance funds in Chile between 1981 and 1994. It describes the dual-track health care system which has emerged in Chile over the course of the last decade, and assesses the implications for equity, access, efficiency, and cost control. The paper then outlines three alternative sets of reforms for the provision of health insurance in Chile.

Over the course of the last decade, efficiency in the provision of services has often been taken to mean a reduction in the role of the central government. Worldwide, the public sector has come under assault for its inefficiency, rent-seeking behavior, and corruption. Demands to restructure the public sector “so that it runs more like a business” have proliferated. This new emphasis runs counter to much of the received wisdom in less developed countries. Indeed, services such as basic education and basic health had been loosely termed “public goods” for decades—and the central government had been regarded as the most efficient provider of both. And yet, encouraged by donor organizations and by some in the academic community, many countries have scaled back the role of the central government in the education and health sectors.

This paper analyzes one such transformation: the creation of private health insurance plans begun in Chile in 1981. This reform of the health sector in Chile was not an isolated occurrence. Rather, the privatization of health insurance was part of an all-encompass-
ing attempt to redefine the fundamental tenets that had guided the Chilean economy for decades. Encouraged by the neoliberal recipes of the so-called “Chicago Boys,” the military junta which ruled Chile between 1973 and 1990 deregulated the economy, freed up prices, slashed tariffs, and seriously cut back the role of the central government in the provision of education, health and nutrition, pensions, and housing.

Reforms in the health sector had two main thrusts. First, the provision of basic health was decentralized: municipalities were made responsible for the delivery of most primary services, while the central government continued to provide the bulk of the revenues (Heyermann 1994). This, it was argued, would bring about improvements in internal efficiency and accountability. Second, the government encouraged the participation of the private sector: private insurance companies, the Instituciones de Salud Previsional (known in Chile by their acronym as ISAPREs), were allowed to collect the 7 percent payroll tax which all Chileans have to pay for health services, and to insure the public directly; large companies could also run private insurance plans limited to their own employees. This, it was believed, would foster competition and further improve efficiency. By the mid-1980s the reforms had “transformed Chile’s health-care system from a centralized public-operated one to a decentralized one allowing for private sector participation” (Castañeda 1992, 68).

This paper focusses on the second aspect of these reforms: the creation of private health insurance plans. The paper is divided into four sections. The first section gives a brief description of the financial flows in the public and private health sectors in Chile as they stand today. The second analyzes the implications of the Chilean system for equity, access to health care, efficiency, and cost containment. The third gives an account of the changes in the health-care market effected by the democratic government of Patricio Aylwin between 1990 and 1994, and assesses the efficacy of these minor reforms. The concluding section outlines two alternative strategies for the provision of health insurance in Chile.

**FINANCIAL FLOWS IN THE HEALTH SECTOR IN CHILE**

Chile operates a dual-track health insurance system. All employed or pensioned Chileans must contribute 7 percent of their payroll toward health insurance for themselves and their dependents. This
payroll tax is collected by a state collecting agency, the Fondo Nacional de Salud (FONASA), which disburses funds to insurers in the private sector, or to health providers in the public sector.

**Private Sector**

As Figure 1 shows, ISAPREs collect funds through three "pipes." First, when consumers enroll in a private insurance plan, FONASA transfers the 7 percent payroll tax to the relevant ISAPRE. This is therefore the effective minimum that an ISAPRE can charge its members. The lion's share of ISAPREs' revenues (77 percent) comes from the payroll tax. ISAPREs can supplement this with additional premiums (10 percent), and a combination of copayments and deductibles (13 percent) (Oyarzo 1992, 10). In addition, ISAPREs receive various explicit and implicit subsidies from the government.

ISAPREs disburse these funds by providing medical services directly with their own doctors, by contracting out services to doctors with whom they have an agreement, or by reimbursing their members for a visit to a doctor of their own choice. ISAPREs

![FIGURE 1](image-url)
vary wildly with respect to the price of their insurance package, the services they offer, and the method of health-care delivery.

**Public Sector**

As Figure 2 shows, FONASA collects money from two pipes to pay for the provision of health for those not enrolled in an ISAPRE: the 7 percent payroll tax (45.4 percent of the total flow of revenues in the public sector), and treasury allocations (another 40 percent). Another 13.9 percent of revenues flows from consumers directly to the providers of medical services in the form of copayments (Oyarzo 1992, 8).

Chileans who do not belong to an ISAPRE may rely directly on the National Health System (NHS), or they may visit a private doctor of their choice under the Preferred Provider System (PPS). In either case, FONASA covers only part of their medical costs. Consumers must pay the rest directly.

The system which establishes copayment levels for the NHS and the PPS is quite complicated. Patients serviced by the NHS are assigned to one of four income brackets: income bracket A (for the indigent), B (for those employed at the minimum wage or less), C (for those who earn less than one and one-half times the minimum wage), and D (those who earn more). For patients in income brackets A and B, there is no copayment; for those in income bracket C, copayment is 25 percent; and for those in income bracket D, copayment is 50 percent (Castañeda 1992, 80–84).

Copayment levels in the PPS also vary—but according to the doctor visited, rather than the patient’s income. All doctors servicing PPS patients in Chile must choose one of three price levels fixed by the Ministry of Health and must then adhere to it for all of their transactions: level 1 (a base), level 2 (the base plus 30 percent), or level 3 (the base plus 60 percent). Patients in the PPS receive a voucher from FONASA before visiting a doctor and use this voucher to pay for part of their medical services. Out-of-pocket payments for a voucher are 50 percent for services performed by a level-1 doctor, 60 percent for a level-2 doctor, and 69 percent for a level-3 doctor (Castañeda 1992, 84–86). By making copayments a function of the physician’s price level, the PPS is set up to discourage the use of more expensive doctors.
The provision of health in Chile is thus highly fragmented. In 1990, 15.1 percent of Chileans were covered by an ISAPRE, 68.7 percent used the NHS, and 12.1 percent used the PPS; the remaining 4.1 percent were covered by special regimes, such as that which insures the armed forces.

**EQUITY, ACCESS, EFFICIENCY, AND COSTS IN THE CHILEAN HEALTH SYSTEM**

**Equity**

Uwe Reinhardt has predicted that the United States will eventually move to a three-tiered health-care system: well-to-do Americans in the top tier will enjoy unlimited access to medical services on a fee-for-service basis; the middle classes in the middle tier will increasingly be enrolled in capitated health maintenance organiza-
tions; and the poor and near-poor in the bottom tier will be relegated to under-funded public hospitals (Reinhardt 1994a). In Chile the method of delivery may be different, but a three-tiered health system is well on its way: the upper and upper-middle classes are enrolled in the ISAPREs; white-collar workers tend to use the Preferred Provider System; and blue-collar workers and the indigent have no choice but to remain in the National Health System.

As Table 1 shows, the ISAPREs have segmented the market by income level. Over 80 percent of Chileans in the lowest two income quintiles, but only 37.2 percent of those in the highest income quintile, were enrolled in the NHS in 1990. Conversely, only 2.5 percent and 5.4 percent of Chileans in the lowest and second-lowest income quintile, respectively, were insured by an ISAPRE, as compared to 44 percent of those in the highest income quintile.

<table>
<thead>
<tr>
<th>Income quintile (by households)</th>
<th>National Health System</th>
<th>Preferred Provider System</th>
<th>ISAPREs</th>
<th>Other</th>
<th>Total number of beneficiaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>84.5%</td>
<td>10.3%</td>
<td>2.5%</td>
<td>2.6%</td>
<td>3,080,425</td>
</tr>
<tr>
<td>2</td>
<td>80.5%</td>
<td>11.1%</td>
<td>5.4%</td>
<td>3.0%</td>
<td>2,857,388</td>
</tr>
<tr>
<td>3</td>
<td>71.3%</td>
<td>13.7%</td>
<td>11.1%</td>
<td>3.9%</td>
<td>2,595,618</td>
</tr>
<tr>
<td>4</td>
<td>58.6%</td>
<td>13.4%</td>
<td>22.3%</td>
<td>5.7%</td>
<td>2,293,510</td>
</tr>
<tr>
<td>5</td>
<td>37.2%</td>
<td>12.7%</td>
<td>44.0%</td>
<td>6.1%</td>
<td>2,083,006</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>68.7%</strong></td>
<td><strong>12.1%</strong></td>
<td><strong>15.1%</strong></td>
<td><strong>4.1%</strong></td>
<td><strong>12,909,947</strong></td>
</tr>
</tbody>
</table>

Source: Ministry of Planning, 1991

It should come as no surprise that, in 1992, 16 percent of the population (those enrolled in an ISAPRE) used 43 percent of the resources spent on health care, while 84 percent (those enrolled in the NHS or the PPS) used 57 percent of the resources. Put differently, each member of an ISAPRE spent four to five times as much on health care as did each member of the public health system (Oyarzo 1992, 35). The private sector in Chile offers quality service, modern technology, and the health-care amenities one might find in a wealthy country. By contrast, the public sector uses outdated, run-down equipment, and provides poor quality service. Public hospitals suffer from serious under-investment: the Chilean Minis-
try of Planning estimates that around 5 percent of the yearly health budget should go toward investment "under normal circumstances"; on average, that figure was barely 2.4 percent in the 1980s. Now, an estimated US $77 million is necessary to bring hospitals in the NHS back to working order (Mideplan 1991a, 66–67).

Health care in the public sector is rationed by monetary and non-monetary means. Queues are prohibitive, and patients may be denied health care outright. Copayment levels have gone up in the NHS—because services were free before 1986, regardless of income—and in the PPS—because fee schedules have not kept up with inflation, and doctors have compensated by moving up price levels (Oyarzo 1991, 70–71); recent government attempts to extract even more copayments from patients in the NHS would be one more step toward stricter rationing. This may all be evidence of "the successful transfer of costs from the public sector to the private sector" (Ardito-Barletta, preface to Castañeda 1992). It is also, however, making it more and more difficult for the poor and the near-poor to receive even a minimum level of health care.

Before the reforms of the 1980s, the Chilean health-care system was redistributive. The mandatory 7 percent payroll tax ensured that the wealthy paid more for their health care than the poor. The creation of the ISAPREs fundamentally changed this relationship: it provided those who "overpaid" with an incentive to leave the public sector, while those who "underpaid" had an incentive to stay. Ironically, what cross-subsidization is now left in the system (other than that represented by the shrinking proportion of FONASA's revenues, which comes from the general government budget) transfers resources from the middle classes to the poor. The Chilean Ministry of Planning has estimated that income brackets C (those who are making less than one and one-half times the minimum wage) and D are net contributors to FONASA; that is, they pay more for health care than they consume (Mideplan 1991a, 66). A system in which the lower-middle classes are evermore burdened with the costs of others' health care, whereas the wealthy are not, can hardly pride itself on being equitable.

**Access**

Access to health care is hard to measure directly. Occasionally, data from surveys in which consumers themselves report on "access" can be used (Freeman et. al. 1987) More often, access is
proxied by a measure of frequency (such as the number of visits to a physician per capita), or a measure of output (such as the health status of the population).

Both the total number of visits to a doctor and the total number of hospital days have risen steadily since the creation of the ISAPREs (Oyarzo 1992, 23–25). But here too there are substantial differences between the private and public sectors. In 1987 ISAPRE members paid an average of 3.73 visits to a physician, as compared to 1.56 for those covered by the public sector. ISAPRE members were also three times as likely to have surgery, and three times as likely to be admitted into a hospital—though the average length of hospital stay was much shorter for those enrolled in an ISAPRE (Jiménez de la Jara 1990, 390); furthermore, these figures do not take into account the quality of the care received, which was surely lower in the public sector than in the private sector.

Numerous indicators show that the health status of the Chilean population improved substantially during the 1980s. Life expectancy at birth increased by ten years between 1970 and 1988; infant mortality fell from 74 deaths per 1,000 live births in 1973 to 20 deaths per 1,000 live births in 1987; and the profile of Chilean mortality changed from that found in most less developed countries, with a preponderance of waterborne and infectious diseases, to that found in industrialized countries, with most deaths due to cardiovascular diseases, cancer, and accidents (Castaneda 1992). It is not clear, however, to what extent these reductions are a result of improved nutrition and sanitation, of targeted programs to improve prenatal and infant health, or of privatization per se. It is, therefore, hard to make a definitive statement about the effect that the creation of the ISAPREs has had on access to health care in Chile.

Efficiency

Privatization of service delivery is often regarded as a way of increasing efficiency. Indeed, much of the rationale for the reforms carried out in Chile was that the health sector suffered from all of the inefficiencies associated with a monopolistic, publicly-operated system: low worker productivity, suboptimal mix of inputs in production, lack of technological innovation, and limited accountability because of the absence of alternative providers. It was thought that these shortcomings would be solved with the introduction of private health insurance plans (Miranda 1993, 8).
The reforms of the 1980s almost certainly improved efficiency on some of these counts. These efficiency gains have been stressed by many observers, and it seems pointless to reiterate them once again. Notably, however, consumer dissatisfaction in 1993 was almost as widespread among those serviced by an ISAPRE as it was among those serviced by the National Health System (Centro de Estudios Públicos 1993). In addition, the reforms introduced a multitude of new flaws into the Chilean health care system.

**Mandatory Payroll Tax**

Once the mandatory 7 percent payroll tax ceased to function as a system of cross-subsidies from rich to poor, and from healthy to ill, it lost much of its raison d'etre. Specifically, the payroll tax has become a source of inefficiency on two counts. First, many consumers in the ISAPRE system are forced to overinsure. Given that they have to disburse 7 percent of their income for health insurance, Chileans opt for the plan that most closely matches their contribution; had they been free to decide how much insurance to buy, consumers might have chosen less. Indeed, the mandatory payroll tax encourages the well-off to insure even against predictable, low-cost medical expenses (such as dental work or eyeglasses), a form of insurance which is generally thought to be inefficient. Second, when an individual enrolls in a private insurance plan FONASA automatically transfers his or her payroll tax to the relevant ISAPRE; if the price of the plan is less than the amount transferred, the ISAPRE keeps the difference (Jiménez de la Jara 1992, 143). Such institutionally protected economic rents have allowed ISAPREs to keep administrative overheads above 20 percent of total costs (Sánchez 1990, 420), while making profits of about 20 percent per year as well—where profits are defined as after tax-earnings divided by capital and reserves (Trabajo de Asesoría Económica al Congreso Nacional 1992). Put differently, the proportion of premiums paid by consumers in the ISAPRE system, which is actually used for the production of medical outputs—the so-called medical loss ratio—has hovered around 70 percent. High administrative costs and persistently high profit levels are both a source of inefficiency, which can only be justified if for-profit health insurance leads to substantial savings elsewhere (Reinhardt 1994b, 117).
Tax-Exempt Status

Contributions by employers and individuals to ISAPRE premiums above and beyond the 7 percent mandatory payroll tax are tax-exempt. This has also encouraged overconsumption of health. Since ISAPRE members are generally in a higher income quintile than those serviced by the public sector, such favored tax treatment is inequitable as well.

Transparency

It has often been pointed out that consumers of health care in a competitive market are faced with a bewildering array of plans with different premiums, different deductibles and copayment rates, and wildly different benefits packages. At best, this is an “information cost” for consumers; at worst, it makes the concept of “consumer choice” meaningless. Most consumers do not even understand much of what is specified in a health insurance contract. This is certainly the case in Chile, where the absence of a standard package of benefits keeps consumers of health care from making rational choices based on reasonably good information. Matters are made worse by the fact that every ISAPRE tries to tailor a particular benefit package to each individual consumer, to allow him or her to spend as close to the 7 percent payroll tax as possible (Sánchez 1990, 425–26).

Lack of Coordination between Sectors

Only recently have ISAPREs been allowed to contract for services with the public sector (Oyarzo 1992). Previously, they had to use their own facilities, or they had to come to agreements with private hospitals or doctors. This deprived the public sector of a valuable source of revenue; it also led to a duplication of structures, and to the underutilization of public hospitals in many parts of the country (Castañeda 1992).

Market Segmentation

Biased risk selection is the most flagrant source of inefficiency in the Chilean health-care system. When the ISAPREs were created, no legislation was written to force them to insure particular individuals or groups. By law, the state was “an insurer of last recourse” (Miranda 1993, 7). This allowed the ISAPREs to segment the market, skimming it for good risks and leaving bad risks to the public sector.
Indeed, since ISAPREs could legally terminate contracts at the end of each year, they often insured people when they were young and healthy, and refused to do so when they were older or chronically ill (Trabajo de Asesoría Económica al Congreso Nacional 1992). This was a particularly galling state of affairs, since those dumped onto the public system had not contributed to the public health coffers during their active years.

In 1990 only 3.4 percent of those enrolled in an ISAPRE were 60 years or older, as compared to 11.2 percent of those enrolled in the public sector. Also, the proportion of dependents per salaried worker or retiree in that same year was much higher in the public than in the private sector; this was due both to different demographic characteristics among income quintiles, and to the fact that ISAPREs often set higher premiums for those with numerous dependents (Mideplan 1991b, 19). Until 1986 most ISAPREs would not even insure women of childbearing age. This changed only when the government decided to pay for maternity leave, and pre- and postpartum care for all women (Castañeda 1992, 101). Since governmental transfers were computed on the basis of income, 62.5 percent of them were given to women in the highest income quintile, and only 1.8 percent to women in the lowest income quintile (Mideplan 1990, 184–86). To make matters worse, these funds came directly out of FONASA's budget, so that only those enrolled in the public health system had paid for them (Mideplan 1991a, 66). Government intervention on this count has therefore been highly regressive.

To summarize, it would be difficult to measure whether the reform of health insurance in Chile in the 1980s led to an efficiency gain or to an efficiency loss. The creation of the ISAPREs did introduce many new sources of inefficiency—many more, perhaps, than had originally been anticipated. This should be a caution to those who hail private health insurance as an unambiguous step toward greater efficiency.

Cost Containment

Containment of health-care costs is much less of a concern in Chile than it is in OECD countries. Some estimates have concluded that health-care expenditures in Chile have remained roughly constant as a proportion of GNP over the course of the last decade (Oyarzo 1991, 60). The World Bank has written, however, that
Norbert R. Schady

"escalating health spending" in Chile could soon "crowd out other sectors of the economy" or "raise the cost of labor" to a point where it threatens the country's international competitiveness (World Bank 1993, 122).

If this is true, cost escalation has certainly been driven by the private sector. Between 1980 and 1990 real per capita costs in the public sector went down. Meanwhile, real per capita costs in the ISAPRE system went up by 27 percent between 1986 and 1990 (Oyarzo 1992, 33). This should hardly be surprising, given that there is very little incentive for the ISAPRE system to contain costs. ISAPREs service the upper echelons of the health-care market, where insurance premiums are mandated to be high by law; closed ISAPREs can reasonably assume that health-care premiums above the 7 percent payroll tax will be passed off to employees in the form of lower wages; deductibles and copayment levels in the ISAPREs are low—as mentioned previously, they account for only 11 percent of the financial flows in the ISAPRE system as a whole; and finally, physicians who have an agreement with an ISAPRE are reimbursed on a fee-for-service basis—this allows physicians to induce demand, and tends to encourage overconsumption of health.

It has often been said that a publicly operated health-care system can contain costs more effectively than a system of private insurance (World Bank 1993, 122). As it stands now, the ISAPRE system in Chile seems singularly incapacitated to deal with the issue of rising costs. It would be most inequitable, however, if the problem of cost control were to fall exclusively on the public sector, simply because it is better equipped to handle it. That would only further aggravate the already alarming quality differentials between the private and public sectors.

**THE 1990s: PERFUNCTORY ADJUSTMENTS**

The government of Patricio Aylwin attempted to resolve problems in the Chilean health-care system with an increase in expenditures on public health (Mideplan 1991a) and with interventions targeted at the sources of inefficiency described above.

**Mandatory Payroll Tax**

The government established that henceforth only 6.1 percent of the payroll tax paid by ISAPRE members would go toward current health insurance; the remaining 0.9 percent would be deposited in
a private fund to provide health coverage upon retirement—when consumers' contributions were expected to be lower and their medical costs higher. These funds would be administered by the Administradoras de Fondos de Pensiones (AFPs), the private pension plans which have operated in Chile since 1981. Also, ISAPREs would no longer keep the difference between the 7 percent payroll tax and the price of the premium. They would have to return this surplus to consumers in one of three ways: by offering better coverage, by allowing consumers to use it toward their copayments, or by investing above and beyond the 0.9 percent in their members' old-age health insurance funds.

Both adjustments have relieved the problem of over-insurance, and are thus steps in the right direction. But there are two matters to consider. First, those who can barely pay for an ISAPRE plan with their full 7 percent contribution might no longer be able to afford it; these consumers—an estimated 200,000 of them—would be returned to the public system (Trabajo de Asesoría Económica al Congreso Nacional 1992). Second, the law stipulates that ISAPRE members who transfer to the public system can take the full amount in their old-age health insurance fund and use it toward their pension plan; these contributions would therefore not have to be used for medical expenses. If, as is likely, consumers were to transfer from the private to the public system when their health costs are higher than their contributions, they would be a net drain on the public system. Under these circumstances, it would make more sense for accumulated savings in the fund to be transferred to FONASA instead.

A more sensible approach would be simply to do away with the 7 percent mandatory payroll tax for those covered by a private insurance plan. This would have to be coupled with a mandated minimum level of insurance—one that does not take into account income levels; otherwise, healthy individuals might not insure at all and fall back on the public sector at times of catastrophic illness. Still, abolishing the 7 percent mandatory payroll tax for ISAPRE members seems a more efficient way of resolving the issue of over-insurance than the reforms of the last few years.

**Transparency**

In 1990 the government created the Superintendencia de ISAPREs, a regulatory body to supervise private insurance plans. The Superintendencia has moved to make the insurance market more
transparent in three ways. First, it has mandated that all ISAPREs provide a basic package of benefits to be specified by the Ministry of Health (Miranda 1993); unfortunately, to date this basic package is very limited and quite ill-defined. Second, to keep ISAPREs from hoodwinking their customers, the Superintendencia has standardized the format of contracts offered by private insurance plans (Jiménez de la Jara 1992, 144–45). Third, the Superintendencia has been instructed to publish information “on any matter of interest to the public” (Miranda 1993, 60); this information includes reports on consumer satisfaction with different ISAPREs.

The creation of the Superintendencia de ISAPREs is doubtlessly a positive development, but it is still beset by two problems. First, it has lacked teeth to enforce many of the regulations with which it is charged; as of yet, there is very little legislation to punish ISAPREs for noncompliance with the Superintendencia’s norms (Trabajo de Asesoría Económica al Congreso Nacional 1992). Second, in the absence of a mandated common-relative-value scale, it will continue to be all but impossible for consumers in the private sector to make well-informed choices. Common-relative-value scales express fees for all medical procedures in terms of the fee of some base unit—such as, say, a set of chest X-rays (Reinhardt 1993b, 15–16). In Chile—as elsewhere where there is private provision of health insurance—the market could be made more transparent if the Superintendencia were to mandate such a common-relative-value scale. Each insurance company would then establish a conversion factor—the monetary value assigned to the base unit. With a common-relative-value scale, Chileans shopping around for private health insurance would have to keep track of only one number for each plan—the conversion factor; they could then use that number and information about quality published by the Superintendencia to make an informed choice to best match their preferences. The Superintendencia could thus “force all providers of health care into highly transparent statistical fishbowls that would be carefully monitored . . . and exhibited . . . to the general public” (Reinhardt 1994b, 111).

Common-relative-value scales have one more advantage: they tend to push prices toward uniformity and could thus be an important tool for cost containment of health expenditures (Reinhardt, 1994b). If this were deemed insufficient, the government could also bracket the permissible conversion factor.
Market Segmentation

The government has enacted minor reforms to encourage ISAPREs to keep their members once they are old or chronically ill (Trabajo de Asesoría Económica al Congreso Nacional 1992); all of these have been ineffective. This should not be surprising: ISAPREs will only stop dumping the old and ill onto the public system when it is very costly to do so. This means that consumers and producers alike have to be nudged toward lifetime insurance.

In Germany, a country which operates a dual-track health insurance system, it is very difficult and very costly to transfer from the private to the semi-public sector. Indeed, this can only be done when an individual lapses into extreme poverty (Reinhardt 1993a). The German example holds a lesson for Chile: if the public system is not to attend to all expensive patients with a shrinking budget, legislation must be passed to prohibit transfers from one system to the other. Of course, consumers would have to be made aware, perhaps by the Superintendencia de ISAPREs, that were they to choose a private insurance plan they would have to remain in the private system for life. Insurance companies worldwide have been extremely reluctant to offer lifetime insurance at a fixed premium—in part because changing technology makes it very difficult to estimate future medical costs. Consumers in the private sector could therefore expect ever-rising insurance premiums in their old age; properly informed, they could weigh relative prices and relative quality of care in the private and public systems and make a decision accordingly.

Alternatives for a More Unified Health Insurance System

The previous section has proposed additional changes to the Chilean health insurance system which would keep the flavor of the current regime. The main feature of this regime is that it offers consumers a choice between the public and the private sectors. Reforms to abolish the 7 percent payroll tax for ISAPRE members, to put in place common-relative-value scales, and to make it all but impossible to transfer from the private to the public system, would simply grease the wheels of two-track health insurance. This would do away with many inefficiencies, but not with the inequity that is perhaps the most salient feature of health care in Chile today.

It is sensible to ask oneself whether such a dispersed system for
the provision of health insurance is better than the alternatives. If it is not, it may be time to reassess the basic tenets of health care in Chile. The 64,000-dollar question then becomes: Should the private or the public sector assume the leading role in the provision of health insurance? In answering this question, policymakers dissatisfied with a dual-track system might reconsider one of two options: a publicly operated, national health system of the sort in place in the United Kingdom; or a privately run, pooled-risk insurance system of the kind which was under consideration in the United States before the 1994 debacle.

The creation of a unified national health system in Chile would mean that all payroll taxes would once again go into a common public health pot. In return, everyone would be eligible for health care in an upgraded version of the current National Health System. Of course, consumers could still visit a private physician of their choice, but they would have to pay for the full amount out-of-pocket. Perhaps, a voucher scheme by which the government reimbursed these patients up to the amount it would cost to provide the same service in the public sector could be put in place—though this arrangement might be administratively cumbersome.

Whatever the economic arguments may be, a return to a mandatory system of public health insurance in Chile is not politically viable. It is unlikely that anyone in the government of President Eduardo Frei other than the minority Socialist Party would invest political capital to do away with private health insurance. Parliamentary opposition to the current centrist, reasonably pro-business governing coalition comes from a right-wing front which would not countenance any such change. This opposition would, no doubt, label any attempt to revive a unified public health system a socialist ploy to return the country to the mayhem of the last year of the Allende presidency. With the military still waiting in the wings, this is one controversy that Chilean democracy cannot afford.

A reform of the health sector which moved toward a pooled-risk, privately run insurance system might be an easier political sell in Chile. Pooled-risk systems can take any one of a number of forms—all of which would share two common features: they would be financed on a capitated basis, and they would make use of group buyers of health care—so-called sponsors (Enthoven 1988). Capitation is important because it turns every hospital or physician from
a "profit center" into a "cost center" (Reinhardt 1994b, 113). Since revenues in a capitated plan are fixed (they are the product of the number of consumers enrolled and the cost of the average premium) while costs are variable (they depend on the amount of health care provided, and on the unit cost of this health care), all profit-maximizing insurance plans have an incentive to limit the amount of health care delivered. Properly channeled, this stimulus could be a powerful vehicle to contain costs. Mandatory use of sponsors is equally important because it keeps insurance companies from segmenting the market by offering every consumer an actuarially fair premium. This could be a reasonable solution to the problem of biased risk selection.

Beyond these common features, privately run pooled-insurance systems can vary considerably. The obligation to purchase a minimum level of health insurance can be placed on the individual—a so-called individual mandate. Risks can then be pooled with the mandatory use of uniform community ratings, or with the framework of "managed competition" which briefly became so popular with American policymakers. Alternatively, the obligation to purchase a minimum level of health insurance can be placed on employers—a so-called employer mandate. Alas, employer mandates are probably not workable in Chile because a substantial part of the population still works in agriculture, in the informal sector, or is unemployed (Portes and Schauffler 1993).

The basis of community rating is simple: different health insurance companies vie for the insurance of a "community" (defined as a small administrative unit, such as a borough, or a specialized district, as is often done with education). The community then selects that benefits package which best matches its preferences, and all members of the community have to purchase the chosen insurance policy.

Despite its intuitive appeal, community rating of this sort has a number of built-in problems. For one thing, it presupposes that there would in fact be more than one insurance company in an area; such an assumption may not be entirely unreasonable for urban Chile, but it will certainly not hold in rural areas. Community rating is also inequitable in the way that decentralization is inequitable: benefits packages would be a reflection of each community's wealth. Poor communities would only be able to pay for very basic
insurance, while rich communities would be able to pay for much more lavish coverage. Moreover, if communities were made up of individuals with different economic backgrounds, the poorest members might not be able to afford the insurance policy which had been selected; these members would have to be subsidized, or they would have to move. Nor is it clear what would happen were an individual to move; unless methods were devised to transfer benefits packages, or to allow individuals to enroll in a plan once an insurance cycle had begun, community rating could serve as a powerful deterrent to worker mobility—with all of the consequences this would have for the smooth functioning of national labor markets. Yet schemes to transfer plans or to allow for mid-cycle enrollment could quickly become a bureaucratic nightmare.

Perhaps the most serious shortcoming of community rating is that it would encourage an unstable, downward-pressing equilibrium in the health-care market. As early as 1956, Charles Tiebout hypothesized that a decentralized system would encourage different jurisdictions to provide different tax-service combinations: some would opt for a "high tax-high service" level, and others for a "low tax-low service" level. Consumers would then "vote with their feet," and settle in that jurisdiction which best matched their preferences. In the aggregate, this exercise in public choice should increase welfare. If, however, services are redistributive (which health insurance inevitably is, transferring resources from the healthy to the chronically ill, and from the rich to the poor), all jurisdictions have an incentive to provide an artificially low level. In this situation, the final equilibrium level of services provided is likely to be lower than it would have been in the absence of perverse incentives. To the extent that this is the case, there is an efficiency loss.

How is this theoretical model relevant to the health insurance market? After all, cynics might argue that a downward-pressing equilibrium is just what is needed to contain health-care costs. Alas, such an unstable, hit-or-miss method to slow down costs hardly seems ideal. Moreover, a high proportion of those who would move to communities which offered generous benefits packages would inevitably be consumers for whom health care costs are highest: that is, the chronically ill. By encouraging this kind of migration, community rating is at best an imperfect solution to the problem of adverse selection.
Managed competition is a more complicated solution which allows individual consumers to choose their preferred health plan. Both the number of potential sponsors and the number of insurance companies offering benefits packages are multiplied. The former can now include large employers, unions, local governments, or a health insurance purchasing cooperative (HIPC) set up expressly for that purpose. In the simplest model, competing insurance plans offer a clearly specified, standard benefits package to all of the residents in an area. Sponsors then collect a premium from every consumer equal to the price of the cheapest insurance plan. Consumers select any one of the plans offered to the sponsor, and pay the difference between the price of the cheapest plan and the price of the plan they have chosen. This monetary disincentive, it is argued, should encourage consumers to enroll in the cheapest plan—and should thus help contain aggregate costs (Reinhardt 1994b, 109–13).

Managed competition is probably an improvement on a system of uniform community rating. Consumers have a good deal of choice—between competing insurance plans and between separately priced perks which could be added to the basic package of benefits. Also, by offering different price-service combinations within a single jurisdiction, managed competition somewhat reduces the incentive to move to follow health-care plans. Yet managed competition does not solve many of the problems inherent in a system of pooled-risk, privately provided insurance. For one thing, it is not as equitable as direct provision of health insurance by the government. Nor is it really practicable where there is not a well-developed, competitive insurance market. Managed competition is therefore no more than a partial solution for a country like Chile: if there is now two-track private and public health care, managed competition would produce a two-track urban and rural health-care system. It is not clear that this would yield a better outcome in terms of equity and efficiency than the revamped dual-track system I proposed in the previous section.

It is, finally, easy to see how the complexity of a system of managed competition could undermine its effectiveness. Indeed, its most vocal supporter concedes that “the complex and subtle tasks that must be carried out by sponsors in a system of managed competition might lead some to conclude that it is all too complicated for either the public or the private sector” (Enthoven 1988, 1978, 1980).
Norbert R. Schady

Administrative constraints could be a crippling problem for a country where managerial capacity is still much more limited than it is in the United States.

CONCLUSION

Whether markets or the state should be charged with the provision of health insurance is a debate that has raged for years. Like many ideologically charged debates, this one has generated more heat than light. Peculiarities of the health care market itself—such as adverse selection, moral hazard, or information asymmetries between producers and consumers—have engendered much confusion. But the discussion has also been plagued by a lack of clarity as to what it is that should or should not be provided by the private and public sectors. The distinction made by Uwe Reinhardt between the two facets of health-care financing is enlightening here: the first facet involves the extraction of funds from private households, while the second facet involves the disbursement of these funds to the providers of health care (Reinhardt 1993a, 44).

When an unregulated private sector is charged with responsibility for the extraction of funds for health insurance, the result is unacceptable in terms of equity. This is because private insurers will always have an incentive to separate good risks from bad risks and to offer actuarially fair premiums. The government must therefore continue to have a leading role in this facet of health care—either by collecting all insurance premiums itself or by regulating insurance funds to ensure that risks are pooled. The case for involving the private sector in the disbursement of funds is much stronger. Competitive pressures in the delivery of health care should lead to an increase in efficiency—though government intervention may still be necessary to mandate a minimum package of benefits and to make the market more transparent. Costs, too, may have to be controlled—with the introduction of budget caps or fee schedules, or with the transformation of three-cornered fee-for-service markets into capitated ones. Creative solutions to these problems can—and have—been found. But it seems that these subtleties were not appreciated by the Chilean government when it first encouraged ISAPREs to enter the health-care market. Or, perhaps, they were all too well appreciated. Cynics might say that a right-wing government had used the chimeric pursuit of efficiency as a Trojan horse: it had helped the wealthy rid themselves
of their responsibility for the well-being of the poor.

Chile has made great economic strides in the last 15 years or so. Its upcoming accession to NAFTA (or, more accurately, AFTA) is a confirmation of this hard-earned progress. Indeed, Chile has consistently kept ahead of the economic curve—first, by realizing the limitations of a state-led, inward-looking growth strategy, and then, by providing safety nets for the most disadvantaged. Not surprisingly, it has become the envy of many of its neighbors. But a greater reliance on the market in the provision of health insurance has not served Chile well.

Notes
1 I have borrowed this framework from Uwe Reinhardt (Reinhardt 1993a).
2 Figures for the public and private sectors are for 1990.
3 In 1983, 49 percent of doctors were registered in level 1, 34 percent in level 2, and 17 percent in level 3; by 1989 these proportions were 9 percent in level 1, 25 percent in level 2, and 66 percent in level 3. For hospitals, a similar, if less dramatic trend can be observed. In 1983, 42 percent were registered in level 1, 46 percent in level 2, and 12 percent in level 3; by 1989 these proportions were 34 percent in level 1, 74 percent in level 2, and 11 percent in level 3.
4 In 1980, 54.5 percent of financing for the public health system came from treasury allocations; by 1989 this proportion had fallen to 35.2 percent (Oyarzo 1992, 22).
5 Alternatively, the system could be set up in such a way that insurance plans offer different premiums to reflect the risk profiles of different groups in the population. Alain Enthoven has argued that this would reduce the incentive for insurance companies to keep the chronically ill from enrolling, and that the sponsor could offset any resulting inequities with the use of cross-subsidies—effectively making the premium that consumers pay to the sponsor uniform (Enthoven 1988). This refinement may seem appealing, but it allows consumers to skim the market for a suitable sponsor: low-risk consumers will prefer sponsors who are less equitable (that is, where premiums are closest to being actuarially fair), while high-risk consumers will prefer those who are more equitable (that is, where premiums are farthest from being actuarially fair).

References


Trabajo de Asesoría Económica al Congreso Nacional (TASC), 30. 1992. Modificaciones a la actual legislación del sistema de ISAPRE.

The author examines the historical and contemporary obstacles to and opportunities for the development of a cohesive national Gypsy identity and argues that, given their marginalized place in European society, Gypsy communities could better position themselves politically and socially by developing a common sense of identity as strong as the undifferentiated, negative perception non-Gypsy societies hold of them. Elements of such a hypothetical program would include cultivation of a high, written form of Romany, development of educational institutions to foster an intellectual elite, the spread of a common history of the Gypsy origins as an Indian people, and psychological identification with some as yet unidentified homeland.

How could a hypothetical, would-be forefather of the Gypsy nation create a unified, cohesive Gypsy identity in Europe today? The question is important for both policy and theory. Europe is home to several million Gypsies, spread across the continent, but particularly concentrated in Eastern Europe. Though socially marginalized, they have a real impact on the larger society, often in ways that the larger society would rather they did not: Gypsies are fantastically
overrepresented in prison populations and unemployment figures, adding to the financial burdens on European society. Moreover, as European birth rates head toward sub-replacement levels, Europe needs to squeeze more production out of the same population. Finding ways to make marginalized groups more active in society will become more important. Since developing a cohesive national identity is one way of bringing Gypsies into the community of Europe, the parameters of such a policy deserve examination.

More importantly, within the Gypsy communities, movement toward a pan-Gypsy identity is already evident: within and between states, cultural exchanges are being formed, schools for Gypsies opened, and conferences held. Among fledgling groups of activists and intellectuals, one senses that a Gypsy identity is attainable and desirable. Without judging whether this is reasonable or not, we may still say that, if Gypsies are to seek a more coherent national identity, they should know what the possibilities and limitations are.

The Gypsy communities are living laboratories for theories of national identity formation. Here is a "primordial group," with many of the characteristics that underlie the development of nations. Here, too, is a present-day experiment for theories focusing on the role of the written word and common myths in the birth of a national people.

**BACKGROUND AND HYPOTHESIS**

Gypsies are a highly disparate group. Culturally, they adhere to dozens of distinct patterns, often corresponding with that of the dominant ethnic group in the region. Language, too, is highly varied: most Gypsies speak the locally dominant language, often as their mother tongue, and cross-group communication is very limited (Fraser 1992, 12).

Critically, there is only a weak sense among the Gypsies that they constitute a single people. Many do not use the word "Gypsy" to refer to themselves as much as to "other" Gypsies—those who are from a different tribe, speak a different language, or have a different traditional occupation.

Nonetheless, shared characteristics exist. A common language, Romany, is found across Europe. Related social patterns have survived even among urbanized Gypsies. Most significantly, the Gypsies—however they see themselves—are despised, mistrusted,
and marginalized by the societies in which they live as Gypsies, as a monolithic entity. This persistent negative identity, is the best rationale for creating a viable, positive Gypsy identity, and the most likely vehicle for developing one.

Thus, the argument is that, given their marginalized place in indifferent or hostile societies that do not distinguish between types of Gypsy undesirables, the Gypsy communities could better position themselves politically and socially by developing a sense of common identity as strong as the perception of commonality already held by the non-Gypsy societies. With common identity comes a common voice, a unified front in seeking whatever it is they might wish to seek. The important point is that whatever they might want, they could achieve it better as a single, stateless nation among nation-states, than as a vaguely identified assortment of disenfranchised individuals and groups.

Because this paper examines a question of what could be, rather than what is, it employs a comparative framework. Lessons are drawn from the development of national identity among groups that shared some characteristics with Gypsies: Jews speaking different languages formed a common identity based on religion and the restoration of Hebrew; nomadic Lapps have articulated a common identity within and across national states.

While no example fully matches the circumstance of the Gypsies, different experiences do suggest lessons for the Gypsies: cultivation of a high form of Romany; development of educational institutions to foster an intellectual elite voice and to spread a common history of the Gypsies origins as an Indian people; and psychological identification with some homeland as a “place of pilgrimage.”

Nationalism, Nation-State, and the National State

A commonly accepted definition of nationalism is “an ideology which holds that the state and the nation ought to be congruent,” thus state and nation are coequal.¹ According to this definition, one observes that there are practically no nation-states in the world, since every state contains citizens from more than one nation, and almost every nation has members in more than one state.

However, this observation is a function of the definition, and a very limiting one because it does not accord with the experience of the senses or logical tests. Is a state not nationalistic if it is not (or
not yet) a nation-state? Must nationalism be successfully implemented in order to be what it purports to be? Catalans agitated for recognition, but settled for strong autonomy in a weak Spanish state. Are Catalans, then, not nationalists? Although the Soviet Union sustained nationalist sentiment in some cases, exacerbated it in others (and, in Central Asia, created it), the Soviet Republics were not states by any stretch of the imagination. Our present definitions suggest that nationalism and the nation-state are epiphenomenal, shadow concepts.

However, if we argue that nationalism is an ideology which holds that the state ought to implement the nation's interests, and thus that the nation-state is a state constituted by (though not solely of) and operated for (though not solely through) the nation, we find that there are a great many nation-states. Most states, in fact, are nation-states, some (like Belgium or Finland) are binational states. By this definition, the attempt and implementation of nationalistic principles stand in closer congruence. Because the term “nation-state” is widely accepted, I will use the term “national state” to define a state that exists for the nation, though it may contain other nations. I will also use “nationalism,” but with the understanding that I do not mean the same thing by it.

This distinction has consequences. If the common definition of the world fits, then other forces such as class, capital, and gender drive state choices, and national questions are epiphenomenal. If the world is more akin to my image of national states, then national identity plays a more meaningful role in the choices of states. Gypsy nationalists could have “room at the margins,” in that space where, by definition, a separate nation can exist within another national state. This is what Gypsies have always done: for a long time that vital margin of territory—the minimum necessary for survival of any human group—was held by Roms, paradoxically, simply by keeping on the move. Gypsies found their living space on the edge of the road, on the wasteland fringes of industrial society, among the rootless of the great cities (Kenrick and Puxon 1972, 214).

**Why not a State for the Gypsies?**

Traditional understandings of nationalism suppose that the nation will attain its own state; it would follow that Gypsy nationalism would also eventually desire a state of its own. There are
obvious obstacles to a Gypsy state. Gypsies are dispersed, and no matter how the map is divided, in no sizable areas do they constitute a majority. All the lands they occupy are already divided into national states. Earlier national movements in Eastern Europe forwarded nationalist claims against imperial entities, but Gypsies would be forwarding a nationalist claim against other national states. Gypsies have missed out on the great European nationalist land rush; staking a claim at the start, hard as that proved, was easier than claim-jumping would be today. For the Gypsies to begin that process today would prove intolerable to those national states who have the good fortune to have put their birth turmoil behind them; simply put, the start-up costs of Gypsy statehood would be too high. Gypsy nationalism could not attain statehood. Is it then nationalism? My definition allows this, since it focuses on the interests of the nation, not on the state for its own sake. With the interests of Gypsies as a nation in view, Gypsy nationalism has attainable goals. I turn now to an examination of the interests of groups which have sought to possess kings without countries, nations without states.

The Idea of the Homeland

If the idea of a Gypsy state must be ruled out, can an effective national identity be created without a homeland? The physical resources of the homeland provide wealth, which translates into power to defend the nation's place among nations and to teach its children its history. In the homeland, generational consistency of culture and values can be assured.

This suggests something far more important than the material advantage of controlling land. A homeland's greatest value is spiritual and psychological. Prideful members of the nation can say: "Here, our way rules; here, things are as they should be, and,"—because no one tells them to do things a different way—"here, we are free." Spiritually, the homeland is a place of refuge.

This preeminence of the spiritual seems the only way to explain why some nations have sought to form states even when doing so is at odds with their economic interests, such as Slovaks in former Czechoslovakia. Nationalist appeal softens the pain of poverty, arguing that a poor but free nation is superior to a wealthy nation controlled by some other, as if to agree with Milton that "it is better to rule in Hell than serve in Heaven."
Heaven and Hell have as much physical substance as some homelands have had. Israel’s location is no coincidence, yet for nineteen hundred years the idea of the Temple, a physical center to Judaism upon which the claim to Israel is based, was kept alive in ritual and story. Generations without experience of Palestine nonetheless knew that the local synagogue was a stand-in, a shadow, because the one Temple belonged on one hill in Jerusalem, though no one they knew had ever seen it. Jewish nationalism was not a developed force throughout the centuries, but the idea of the homeland kept a disparate people united enough to allow that sense to be built upon by later nationalist movements. Before 1140 Judah Halevi, a Spanish Jew, wrote of Jerusalem:

Beautiful heights,
Joy of the world,
City of a great king,
For you my soul yearns,
From the lands of the west.

My pity collects and is roused
When I remember the past:
Your story in exile,
Your temple destroyed.

I shall cherish your stones and kiss them,
And your earth will be sweeter
Than honey to my taste (Gilbert 1978, 11).

It is reasonable to ask how necessary full sovereignty is. How much must the nation be in possession of its homeland to enjoy its benefits, or to enable a nationalizing elite to mobilize its people around the idea of that homeland? Must the territory function with limited sovereignty?

Recall my definition of nationalism: the desire of a nation to wield effective control over a territory. It follows that other nations can find some space within a national state, since equilibrium is possible at less than all-or-nothing. A nation could form around the idea of a homeland without impinging on the claimed territory of another nation.
A GYPSY HOMELAND?

This conclusion has obvious consequences for Gypsy nationalist policy. The mere identification of a homeland, even India, might have a "catalytic effect," focusing disparate cultural groups. A city such as Skopje might be declared the "capital" of European Gypsies, serving as a base for congresses and research. Even better would be cultural autonomy within some European state, such as Macedonia.

The Effects of Regionalism

Some obstacles to this goal might be mitigated in the future. The integration of the European Union (EU) has important consequences for sovereignty and the national state. The states of the EU will not disappear, but it seems increasingly plausible that they will cede important elements of sovereignty.

The EU recognizes various historical and political regions of Europe, many of which are linguistically or ethnically based. These regions have a representative body and a degree of legal autonomy within the EU, if not always at home. In time the regions may become means of securing representation and access to power. The states will not fade away, but weaken somewhat, and the regions will become political forces in their own right.

In this scenario there is an opening for Gypsy nationalism. Of course, these regions need not show any more willingness to grant cultural autonomy to Gypsies than states have. However, because these regions are not homogeneous—because they contain "non-coregionalists"—some provision would be made for "cross-transferability of rights" between regions. It is difficult to imagine that Paris would agree to an autonomous Brittany actually restricting the rights of Frenchmen there. Instead, citizens of a region would retain certain rights even in another region. The region would be the legal point of reference, but the ethnicity that underlies that regionalism would be the rationale.

Such a formula would reconceptualize the interplay between citizenship and territoriality. Rights would still stop at the frontiers of the EU, as they do today; however, within the EU, there would be an "unbundling of territory" with overlapping political allegiances, in something of a return to the organization of the Middle Ages. Many of these unbundled allegiances may be based on national or ethnic claims. As Hasner cautions,
the great achievement of the modern state... the primacy of common citizenship over privilege based on bloodlines... is being jeopardized both by the trend towards transnational de-territorialization and by the trend towards national, subnational, and ethnic reterritorialization.¹⁰

In this context, Gypsy activists could claim a right to representation, not on a territorial, but on a national basis, since the territorial nations—the regions—would in effect have carved out non-territorial rights for themselves. Gypsies could claim the *fruits* or territoriality: cultural protection, educational autonomy, ethnic voting rights. Although consistency is not a hallmark of real political processes, it might well be difficult—more difficult than at present—for a regionalizing Europe to refuse such a claim.

**Voting Rights and Ethnicity**

Nonterritorial rights might include voting rights based on ethnic affiliation; these are in fact increasingly in use in Europe today. Consider the Lapps, or Sami, who live in an area divided between Norway, Sweden, Finland, and Russia, and are a minority even within Lapland. Many Lapps are nomadic, and their traditional migrations cross state boundaries.¹¹ Treatment of Lapps, though not as harsh as that of Gypsies, was hardly benign: forced integration was policy and practice, especially in education, until World War II. Until the 20th century, the few Lappish books were missionary texts, and most Lapps were illiterate.

In the late 19th century, Lapps showed signs of nationalist development: they studied at university, published secular books, and formed cultural clubs. Lapp dialects were standardized in 1913. Calls for political reform followed, but Lapps never seriously agitated for a state. Instead, Lapps carved out a political and cultural sphere with self-contained units in each state, which were also integrated across states.

The intrastate sphere includes institutions supported by the states. Norway’s Sami Parliament has limited cultural authority and consultative “first rights” on issues affecting Lapps. Norway is divided into 13 territorial constituencies in which only Lapps vote. Thus Lapps have ethnic voting rights even outside their homeland. The Finnish Lapp Parliament has only consultative rights, but since 1991 it too has extended voting to Lapps in the south.

At the interstate level, a Nordic Sami Council coordinates the
state bodies and consults the governments. How much these institutions have fostered a common Lapp identity, or are the product of one, merits investigation. We may speculate, however, that these institutions create incentives for individual Lapps to identify with the larger community, because real electoral access is available to those who do.\textsuperscript{12}

Although the Lapp successes have relied on the willingness of the Scandinavian states to develop an unbundled system, they do prove that autonomy is possible within the confines of the national state system. If the Gypsies cannot obtain an analogous system, it is because of political and historical \textit{circumstance}, not because such a system is impossible. The best evidence is the progress Gypsies have already made: Hungary has recently revised its laws on national minorities, giving Gypsies ethnic voting rights.\textsuperscript{13}

In communist Hungary, Gypsies were an “ethnic minority,” whereas others such as Germans and Slovaks were “national minorities” with greater rights.\textsuperscript{14} The new minorities law of 1993 seeks to reconcile a civic definition of citizenship with assurances that minorities can participate either as ‘regular’ citizens or as members of an ethnic group. The law defines minorities as: any group with at least 100 years of residence in Hungary and with its own language and culture . . . \textit{[T]he choice of identity is voluntary. . . . It declares that minorities have a right to set up their own cultural and heritage organizations including local and national self-governments, ensuring their cultural autonomy . . . [A] national and ethnic minority fund be set up within a year to assist the minority self-governments. (RFE/RL Daily Report, 8 July 1993, 6)}

Of interest here are the local and national self-governments with state funding and authority over cultural and educational affairs. These self-governments have a territorial basis, overlapping with other local governments for majority citizens. The new law also adds 13 seats for minorities to Parliament, for which the minority must field a national list.\textsuperscript{15}

In this system there are clear incentives for Gypsies to develop a more cohesive identity. The local self-government system, based on territorial units, does not distinguish between types of Gypsies, encouraging cooperation between different groups. The parliamentary changes provide even stronger incentives for cohesion. With only one list, all candidates have to compete for all votes. Inevitably, compromises will have to be made, encouraging the
political integration of the various groups, which would come to recognize that their interests were served, not by Beás, Romungro or Olah, but by the whole Gypsy community.

The Hungarian voting system, though new and untested, suggests a pattern that Gypsy groups are pursuing: the creation of common identity at the state level as an essential element of the pan-European movement. They have followed this pattern before.

**Gypsy National Identity: The Historical Record**

In addressing the possibilities for a rising Gypsy nationalism in Europe, it is useful to ask if Gypsies have demonstrated national sentiment in the past. If so, we may speak, not of a rising, but a resurgent nationalism. We are not looking for the lost kingdom, merely signs of life before our day. Such signs—though few and tentative—do exist.

Miloslav Hroch identifies three stages in the growth of nationalist movements: cultural and literary activity; pioneering militancy and politicization of the elite; and mass mobilization. Gypsy "nationalism" in the 19th and early 20th centuries engaged in considerable activity characteristic of the first stage, some—more sporadically—of the second, but none of the third. There was a Hrochian retrogression among the Gypsies at mid-century, such that today they are back in the first stage, tentatively pushing toward the second.

**First Contacts**

The origins of the Gypsies in India have been established by scholars, but Gypsies themselves had lost that knowledge. Nomadism did not encourage retention of a common cultural heritage. Rather, Gypsies adopted elements from the cultures whose territories they passed through. Thus the first record of Romany in England, published in 1547 as an example of "Egipt speche," includes words borrowed from Greek and Romanian.

The arrival of the Gypsies provoked a strong reaction. Their first appearance in Germany, at Hildesheim, was recorded in 1407; the first anti-Gypsy law was promulgated in 1416 (Hancock 1991a, 11). Little attention was paid to variations among Gypsies, who were mistaken for Egyptians or Turks. "[E]asily identifiable as a 'foreign element' . . . [they] were quickly singled out as enemies of the state, possibly spies of the much-feared Ottoman Empire." Thus Euro-
peans viewed them as threatening, and, significantly, as alien and undifferentiated. A negative identity was established at first contact.

Such identities need not be exclusively negative, if a group can establish itself, by virtue of its alien identity, as the sole provider of some service. Gypsies occupied such a position, thanks to their specialization in certain trades. The benefits for group cohesion did not always outweigh the costs of the method employed, however, because Gypsies occupied a very particular niche indeed. Beginning in the 13th century, they were held in slavery in the Romanian provinces. General emancipation was not completed until 1864 (Crowe 1991, 66–67).

Crowe speaks of the “social death” of Gypsies during enslavement, and indeed, the identity of Gypsies was reduced to a shadow or negation: Gypsies were what Romanian peasants were not. The movement to emancipate Romanian peasants was cast in contradiction to the Gypsies: “landlords have been in the habit of selling serfs (vecini) like slaves (robi), treating them like Gypsies (Crowe 1991, 61–64).” The burden this endemic attitude would place on later Gypsy organizers should not be underestimated.

Until emancipation, Gypsies had no collective opportunities to undertake the process of cultural activity, political agitation, and mobilization that Hroch describes. Thus timing was critical. Gypsies were emancipated after the last great national claims had been declared. They had spent the critical centuries of European nation-building in the social death of slavery.

**Emancipation**

With emancipation, Gypsies had greater possibilities for organization. In 1879 a conference was convened in Kisfalu, Hungary, “to discuss ways to achieve Gypsy political and civil rights throughout the Continent. Though ridiculed in the press, it stimulated a greater sense of ethnic identity.” (Crowe 1991, 68)

In the interwar period, the first signs of a Hrochian third stage appeared: mass participation. An All-Russian Union of Gypsies ran 80 collectives. Calls for a Gypsy homeland appeared (though it is charitable to say they were not given full attention); these showed influences from Zionism’s calls for a Nachtasyl in Uganda or Argentina (Halperin 1969, 128 and 154).

Romania saw the greatest Gypsy activity. Gypsy journals were
published—Neamul Tiganesc (The Gypsy Family), Glasul Romilor (The Voice of the Roma) and O Rom—and the short-lived General Association of Roma in Romania advocated a national holiday honoring Gypsy emancipation, as well as “a library, a hospital, and a university for Rom.” The association convened a Gypsy World Congress in Bucharest in October 1933. It “opened with the theme ‘United Gypsies of Europe’, [and] adopted a program that advocated efforts to stimulate Gypsy ethnic consciousness in league with greater demands for national minority rights” (Crowe 1991, 69).

A pattern appeared in the interwar period that has been revived in the postwar era: a two-tiered approach, with state-level and interstate-level organization. The more significant activity was in organizations representing Gypsies within states. The many calls for a homeland by the “royal” Kwiek family in Poland had less impact than the coalition of diverse Gypsies who identified with this otherwise dead-end idea. Modifying Hroch, the Gypsies set about attaining national cohesion within each national state, as a prelude to and engine of the ‘international national movement’, Pan-European Gypsyism.

**Collapse**

By the 1930s, real progress was being made toward a Gypsy consciousness. However, the rise of the Fascist states and the Porajmos, the genocide against the Gypsies, snuffed this movement out. The anti-Gypsy reaction developed in tandem with the Gypsy movement. A conference on “The Gypsy Filth” (Der Zigeunerunrat) held in Swabia in the late 19th century proposed a system “whereby bells would be rung in villages as a means of signaling [the Gypsies'] presence” (Hancock 1987, 103). By the 1930s, strict regulation was imposed on Gypsies, and increasingly imaginative proposals exhibited a perverse symbiosis with the Gypsies own initiatives: about the time the Kwieks were petitioning Mussolini for a homeland, Le Temps reported, under the rubric “Gypsy Island,” that “the mayors [in Oberwarth, in Burgenland] have endorsed a proposal . . . to examine the establishment of a Gypsy colony on one of the Polynesian islands.”

By the end of the 1930s, the Nazis abandoned such notions and embarked on the extermination of the Gypsies as the only ethnicity besides the Jews slated for immediate destruction. Estimates place
the Gypsy victims of Nazi genocide at 200,000 to 500,000; perhaps one quarter to one half of Europe’s Gypsies were executed. Gypsy national movements were devastated; the few Gypsy intellectuals were wiped out. Gypsies in Germany were afraid to come forward because pre-Nazi laws still in effect would have consigned them to detention centers (Liegeios 1986, 93). The Nuremberg trials paid no attention to the Gypsies; reparations were not offered.

**Revival**

In the West the Gypsy national movement revived and pursued its agenda after the war. Western Europe was more disposed to hear such appeals after the excesses of the war, though resistance and indifference persisted. The World Romani Congress succeeded the defunct Bucharest Congress (Kenrick and Puxon 1972, 211–212). The organization persuaded the Indian government to acknowledge Gypsies’ Indian origins, thus establishing a claim to a homeland. Gypsies obtained permanent consultative status at the UN in 1979.

Eastern European Gypsies were frozen out of such movements by the communist states. Despite the resurgence of anti-Gypsy rhetoric since 1989, the fall of communism is a positive development for Gypsy nationalists, who now are free to organize in the areas of greatest Gypsy concentration.

Perhaps it is simply too dangerous to organize the Gypsies at all. The history of reaction is a cautionary tale for would-be nationalists: if Gypsy unity is seen as encroaching on the rights of established nations, strong, even hostile resistance may result. Gypsies, having missed out on the chance to stake a claim in the nation-state land rush, might be better off lying low, in a cultural and political sense, than risking a showdown with powerful proprietors.

This caution is not an ultimate objection, however, for the situation in Europe has changed. The reaction to Gypsy nationalism was only part of a larger wave of aggressive fascism which is now discredited; it was not a direct reaction per se. Moreover, most European states have had relatively stable borders since the war. The changes since 1989 have been conducted along recognized internal and external borders, except in Bosnia. The national states of Europe, or at least the short list of contenders, have been decided, and those states are showing an increasing proclivity for regional cooperation that should mitigate their more aggressive
tendencies toward minorities. If this is true, we may expect that this time a rising Gypsy nationalism, while unwelcome, would not meet with so harsh a reaction.

**THE ROLE OF LANGUAGE**

Mihail Kogalniceanu, the emancipator of the Gypsies, predicted that the Romany language would die out, since “in becoming civilized, [Gypsies] will experience new concepts, and not retain so defective a language” (Hancock 1987, 47). Yet, despite strong linguistic assimilation, the “defective language” survives as one of the most important cleavages among Gypsies. There are over 60 major dialects of Romany and group distinctions are often related to language differences. Perhaps half of all Gypsies speak a European language instead of Romany.

Yet the modernist conception that language may create a nation where none exists provides an avenue for developing a unified, if not unitary Gypsy nation. By consciously creating a standard literary dialect, other nationalists have created linguistic uniformity where none previously existed. This allows the transmission of cultural symbols and the sharing of values, which build a community, not merely of language, but of myth and symbol. Let us look at the example of Hebrew.

**The Revival of Hebrew**

The Zionist movement relied in no small part on the revival of spoken Hebrew, which was a “dead language” much like Latin. Indeed, the spoken language was essential to the creation of a unified body of Jews, not only able to communicate, but also identified and identifiable through their language. A Jewish language was needed.

Yiddish or Ladino might have served. Both were closely identified with Jews, and had wide currency. Both were living languages, adapted to modern life. However, both had restricted appeal since they were already identified with a particular subset of Jews: Yiddish with Mitteleuropa, and Ladino with the Spanish diaspora. These languages already served the purposes for which Hebrew was ultimately ordained: they defined a people. Their definitions were too narrow because they did not include all Jewish people—the peoples they would have formed would have been Yiddish or Ladino Jews, not simply Jews. Because Zionism sought to build a
single Jewish consciousness out of many Jewish identities, it needed an ethnic lingua franca not tied to any one of them, though present in all.

The success of that linguistic program is well known. Hebrew is the language of Israel, and Jews throughout the world consider familiarity with Hebrew an important part of their identity. Given that it contributed to the unification of worldwide Jewry and thus the creation of Israel, the resurrection of Hebrew may be Zionism’s greatest achievement.

Prospects for Romany

The program for Gypsies would seem clear, in light of the successes of other created and remembered languages. The Sanskrit-related Romany “language-field” is the obvious candidate for nation-imagining language, a pater linguisticus. Familiar to all the groups, it is not the sole linguistic resource of any. All Gypsy groups speak some other language too, and thus Romany would not be identified with any one group.

A standardized version of Romany would have to be promulgated by combining elements from several fields or dialects into a compound version. New vocabulary would have to be invented. Most importantly, a standardized version might disproportionately benefit certain Gypsies: Eastern over Western, literate over illiterate, settled over nomadic. These points have not escaped Gypsy activists. Consider this assessment of the linguistic agenda presented at the Third World Romani Congress in Göttingen in 1981:

It cannot seriously be believed that the unification movement is unaware of the divisions now separating the Romani population. . . . This, and standardization of the language is of vital concern to members of the World Romani Congress. Aspects of this . . . may be summarized thus:

1. Because of historical and contemporary factors . . . which continue to divide the Romani-speaking populations, there are today a great many widely differing dialects. . . .

2. Perhaps the greatest obstacle in achieving political and cultural unity is the lack of communication among the various Romani groups. . . .

3. [P]rogress toward reunification would be more easily made if a common dialect were available to all groups. To quote Sarkon, e dobinda la shibake ketanemaske s'o paso o angluno karin le
ketanimastes arekh nyamo, the achievement of linguistic unity is the first step toward unity as a people.

4. No single dialect spoken anywhere is so close to the common protoform that it may be adopted without modification. In other words, whatever dialect will be chosen will have to be adapted to a more internationally acceptable form, especially lexically.

5. The propagation of such a standard will be very unevenly achieved. Settled, already literate Rom, such as predominate in Eastern European countries, will have a far better opportunity to acquire the standard dialect. For illiterate and nomadic Rom the task would be much harder.

6. Not all Rom everywhere will ever learn, or be disposed to learn, such a dialect. This will create a 'linguistic elite consisting of those who have learned the new international standard' (Hancock 1981, 15–16, emphasis original).

NEGATIVE IDENTITY AS BURDEN AND TOOL

Two important questions arise: What would vastly different Gypsy groups gain from a common identity, and why should they not want to maintain their separate identities? Gypsies themselves are very aware of their differences—sometimes to the point of referring to themselves as a particular tribe, and to all others as “Gypsies.” Yet there is one sense in which Gypsies are indeed very much alike. Gypsies—however they consider themselves—are uniformly despised, mistrusted, and marginalized by the societies in which they live as Gypsies, as a monolithic entity.

Most Europeans, though vaguely aware that Gypsies have tribes, do not distinguish among Gypsies. Of course, this is also true for other groups. An Englishman speaks of “Italians,” but barely distinguishes between northerners and southerners, while an Italian can give an exacting typology of Italians by region and temperament—Sicilians, Calabrese, Piedmontese—yet knows only that “Englishmen” are “cold and stiff.”

The difference is that, though Italians and the Englishmen know their own tribes diversity, they too share in this undifferentiated vision of oneness. They see themselves as Sicilian and Italian, Liverpudlian and British, because they remind themselves—in sport, school, in reading newspapers, and in voting—of their larger identity. The territorial state is the key. It constantly reminds its people, didactically and by example, who they are. Moreover, that
state can remind other states of who its people are. Through trade and propaganda, Italy creates Europe’s image of Italy, in accord with its self-image.28

The Gypsies lack these resources, and so when they are lumped together by the rest of Europe, their practical identity—the one that interacts with the outside world—becomes wholly ascriptive and received. Their own images of themselves have little relevance to the world, which treats them as a monolith. Europe views them as one, and because Europe does not like what it sees (or thinks it sees), it treats that one Gypsy as a threat, a pariah.

How does Europe even know it is looking at a Gypsy? Negative identity rests upon theories of primordialism and ethnic markers. For example, some groups are less likely to be assimilated due to skin color. The Gypsies, upon arrival from India, looked quite different from the natives; it was easy to identify them. Some intermarriage and assimilation has occurred, and many Gypsies can pass as white—doubtless even more white people are blissfully unaware of their Gypsy genealogy. On the whole, however, markers such as skin color and facial features serve remarkably well to identify a group—and isolate it. Where the desire to isolate it breaks down, assimilation can be rapid.29

Where it holds, individuals may assimilate, but the group will remain identified, as a whole, by the majority. A similar process may be identified in the relations of blacks and Native American to the white majority. Whites ignored the differences among blacks from different parts of Africa during slavery, and summed up their appreciation for the myriad Native American cultures with the notion that the “only good Indian is a dead Indian,” leaving little room for differentiation on this side of the grave.

**Pariahs into Gypsies**

Thus a negative or ascribed ethnic identity can be asserted through such markers. It is also the vehicle by which a broader identity could be further developed within the Gypsy community: a conscious recognition that the outside world thinks of and treats the many Gypsy groups as one. These groups share a common fate, and have suffered a common lot. There is no better recipe for group cohesion.

Blacks from dozens of cultures formed, under extreme pressures, a much more homogeneous group of American blacks—the
common bond among these groups was their common suffering and their common, ascribed, negative identity. Native Americans, too, have adopted a dual-level identity. Many have strong affiliations to their tribe and, as the American Indian Movement (AIM) demonstrated, a sense of common identity based upon common experience. Blacks and Native Americans were treated as unitary entities; they rationally responded by creating an image of themselves that mirrored what had been ascribed to them.

What would a Gypsy nationalist have to do to create a more active identity? The most obvious course would be to foster this sense of common suffering. The most compelling example of that suffering is the Porajmos. This is what the World Romani Council has emphasized: one of its five main sub-bodies deals with Holocaust reparations. In Hungary, a memorial to the Gypsy Holocaust victims was recently erected. Gypsies have also sought, and eventually received, representation on the U.S. Holocaust Memorial Council.

Representation and reparations have value beyond the petty power they bestow on Gypsy functionaries; they establish a universal claim of suffering around which a people can form. This message has not been lost on other groups; it in part explains the hostile reaction of some Jewish groups to Gypsy demands for representation on the Holocaust Council, since the symbol of suffering, and the homeland, might be vitiated if shared. These groups have recognized the connection between love of the homeland and the bond of suffering, the greater power of grief in the national memory, as Renan put it—a power greater than triumph. Since 1945 the Jewish homeland has been Auschwitz as much as Israel.

It is the symbolic value of suffering that is most important. Gypsies lack a territorial base from which to operate, but their suffering may have produced a homeland, the camps, and a common history, the Porajmos.

 Alternatives

It is possible that efforts at a common unifying identity could have a negative effect by threatening the diversity of present identities. The cultural and material gains arising from a unified identity might be offset by losses of diversity. There is no full rebuttal to this objection; some loss is inevitable because change is
inherent in the process of identity unification. However, the Gypsy case may suffer less from a loss of diversity than other national groups, because a Gypsy identity would not be predicated on accession to real territory, and consequently would not make the same demands for homogeneity of the populace required by state-building nationalisms. Where no state is to be built, uniformity need not be complete. The Gypsy identity could be weaker than the identity of a nation-state, and therefore more tolerant of regional variations.

Notes

1 Gellner (1993, 3), for example, shows how nationalism can be benign and even beneficial, and nonetheless follows this definition closely, concluding that “the problem of nationalism does not arise for stateless societies.”

2 Breuilly defines nationalism as a political movement seeking or exercising state power and justifying its actions by asserting that the nation has an explicit character, that the interests and values of the nation should have priority over all other interests, and that the nation should be as sovereign as possible.

3 This actually happened: postwar proposals for a Romanestan, “a territory that would serve as a refuge in the event of persecution” were forwarded by Ionel Rotaru, and in Tito’s Yugoslavia an autonomous region in Macedonia was considered. Earlier proposals called for a Gypsy state on the Ganges, in South Africa, or in Italian Abyssinia. Nothing ever came of them, and it is doubtful they were ever seriously considered.

4 See Gellner (1993): “The infrastructural investment made in [high culture] can be relied on to perpetuate [those cultures]” (p. 122).

5 Renan’s observation that “where national memories are concerned, griefs are of more value than triumphs, for they impose duties, and require a common effort” suggests that the lack of a homeland may be more powerful for unity than its possession: a nation seeking to regain a territory draws its strength from the idea, not from the real estate. Gypsies have never had any land to lose, but if they could be made to feel as if they had, as Zionists felt about Jerusalem, the national principle could operate.

6 Some actually saw it: Jewish settlements remained, and at the end of Crusader rule, 300 rabbis from Flanders and Provence travelled to Palestine to help build up the Jewish communities (Gilbert 1978, 12).

7 Gellner (1993, 107) also notes that “in the diaspora, the Jewish religion referred back to Jerusalem,” adding that this reference had no territorial base to back it up, and thus was a reference to the idea alone.

8 The phrase “and that territory should encompass all the nation,” is the corollary, usually applicable, but not always. Some nationalists may be
satisfied with less. Such people exist; present definitions of nationalism do not have a place for them.

Ruggie (1993) calls the EU a layered, “multiperspectived polity” (pp. 171–172), and gives examples of alternatives to the territorial state system that need not be territorial (extensions of kinship rule), fixed territories (nomadic control of rights of passage), or exclusive (feudalism’s overlapping sovereignties).

See Hasner (1993, p. 53). For ethnic groups that lost the nation-state land lottery, jeopardizing that achievement through deterritorialization is another way of saying “second chance.”

Some Lapps have always been sedentary, highlighting both a difference between Gypsies and Lapps, because Lapps can claim a territorial base, and a similarity, because the Lapps are several different peoples, who mostly share a ‘common plight.

Horowitz (1985, chap. 2) discusses the utility of ethnic affiliations.

The logic being that Gypsies have no homeland or mother country.

The first mandate requires 3,000 votes and each additional mandate 30,000.

Noted in Hobsbawm 1990, p. 12. I adopt Hroch’s idea to my stated definition of nationalism.

Hence the name. The other common form—Zigeuner, cigany, tsiganes—derives from the Greek athinganoi, a heretical sect, a misidentification with less than positive connotations.

Guglielmo 1992, p. 50. Their nomadism did not endear them to the peasantry or townsfolk either.

Jews and Armenians dealt in money, not simply because of proscriptions against usury, but because they were culturally different—they isolated the “oath-dissolving contagion” that money represented in a barter system. They had an identification imposed from without, but one that efficiently maintained the unity of the group. See Navari 1981, p. 26; and Gellner 1993, p. 103.


In 1938, the National Socialist Party declared the Gypsy question to be “mit Bestimmtheit eine Frage der Rasse.” A certain number of Gypsy families were to be spared and kept in a compound, for study by anthropologists. Kenrick and Puxon (1972, pp. 183–184) estimate nineteen thousand deaths, while Huttenbach (1991, p. 45) estimates five hundred thousand. See Huttenbach (1991) for surveys of the Porajmos; see László Karsai (1992) for a study of the Holocaust in Hungary.
Linner (1984) discusses the relationship between Hebrew, the Hebrew press and Jewish nationalism in the context of Jewish Messianism.

See Anderson (1991, pp. 84-85) for an indirect example on the development of German as a “particular-national” language, which selected out German speakers as a discrete and favored group.

The actual spread of a national language is not as important as its psychological value: Nynorsk, an official language of Norway, is a consciously created compound of several dialects. Italy became a national state even though a majority of its citizens did not report Florentine Italian as their native dialect or language until the 1970s.

See Gellner (1993, pp. 44–48) on the tendency toward a unitary form of the language.

“[O]fficials felt that . . . [Gypsies] had neither a protective state, a history, nor a culture and civilization through written languages to justify their claims to minority rights” (Crowe 1991, p. 69).

The catch-22 of Gypsy nationalism is the relationship between education and assimilation: Gypsies have ironically avoided assimilation through their very marginalization. When they have had the opportunity, however, they have demonstrated some tendency to assimilate, and better educated Gypsies do assimilate more readily than uneducated ones (Crowe 1991).

Kenrick and Puxon (1972, p. 214) say Gypsies were “the first Blacks in Europe, and they are the last to raise their standard and seek emancipation.” Descendants of blacks with various African identities forged a more homogeneously identity under pressure.

References


Kings without Countries


The paper uses a simple theoretical model of potential parents’ contraceptive decisions to analyze the effects of different child-support laws on contraceptive care, the probability of accidental pregnancy, and the distribution of welfare. Among other things, the model implies that under noncooperative contraceptive decision making the rules of paternal, maternal, and joint liability are all inefficient, while a combination of a liability rule and a tax is efficient. Under cooperative contraceptive decision making, it shows that the effects of these rules are reversed: the combination of liability rule and tax is inefficient, while the others are all efficient.

Governments sometimes impose child-care obligations on absent parents. In particular they may require an absent parent to make payments to the caregiving parent or to the government. They may do so in part because they want to distribute the costs of child care in a way they consider fair. But they may also be interested in the effects of child-support laws on the behavior of parents and potential parents. For example, they may want to encourage mothers and fathers to stay with their families because of a concern about the effects of single-parent families on children’s welfare or in order to reduce the cost of grants to poor single parents. Finally, they may want to influence people’s decisions about how much care to take to avoid conceiving unwanted children.

In analyzing these issues, it will help if we distinguish two cases that may be of interest to policymakers. First, we could consider the contraceptive decisions of potential parents who do not want to

*M.P.A. Candidate, Woodrow Wilson School, Princeton University; M.A., University of Otago; B.A., University of Auckland.
have a child. In this case, children are accidental and unwanted by-products of sex, and the policy issue is how to induce the potential parents to take due care to avoid becoming actual parents. Second, we could consider the child-care decisions of actual parents. In this case, the policy issue is, Who should look after the child and who should pay the costs of that care?

This note offers an economic analysis of the first of these questions. It attempts to elucidate some of the issues involved in the choice of child-support policy by examining the implications of different legal rules for contraceptive choices and the probability of childbirth. The paper will assume that the individuals involved are perfectly rational and perfectly informed. It should thus be considered as an initial step in the analysis of policy; a fuller analysis would consider the implications of dropping the assumptions of rationality and perfect information.

**A Simple Model of Contraceptive Care**

When children are unwanted and accidental, we can usefully apply Shavell's (1987) model of accident law to child-support law. Accidental pregnancies are different from the typical accidents of tort law in that the parties involved are typically not strangers. But there are enough similarities to make Shavell's analysis interesting. Applied to accidental children, the model is as follows.

Given that a couple have chosen to have sex, they need to weigh the costs of taking different levels of contraceptive care against the potential costs of having a child. The costs of taking contraceptive care are the costs of obtaining contraceptives (money, time, and embarrassment, for example) as well as any costs associated with using them (less sexual pleasure and health risks, for example). The probability of having a child is a function of the level of contraceptive care taken by both the male and the female sexual partner.

The loss borne by the couple, $CL$, is given by the following equation:

$$CL = x + y + p(x, y) L$$

where

- $x \geq 0$ and $y \geq 0$, and
- $x = \text{the level of contraceptive care taken by the man}$,
- $y = \text{the level of contraceptive care taken by the woman}$,
\( p(x,y) \) = the probability of pregnancy as a function of \( x \) and \( y \), and
\( L = \) the cost of having and taking care of a child.

\( L \) should be interpreted to include any cost of a child, including for the woman any cost of being pregnant. We have implicitly assumed that the per-unit cost of both \( x \) and \( y \) is one. The analysis will not consider the effects of price changes on outcomes of interest, so there is no harm in this assumption. We assume further that \( p_x(x, y) < 0, p_x^2(x, y) > 0, p_y(x, y) < 0, p_y^2(x, y) > 0, \) and \( p_{xy}(x, y) > 0 \), where the subscripts denote partial derivatives. That is, increased care by either the man or the woman reduces the probability of conception, but at a diminishing rate, and that increased care has a bigger effect the lower the other partner’s level of care. Finally we assume, for the moment, that the optimal contraceptive choice involves both the man and the woman taking some care.

The couple’s joint welfare is maximized by minimizing the expected loss from having sex, which implies that the optimal level of male contraceptive care, \( x^* \), satisfies
\[
p_x(x, y) L = -1
\]
and \( y^* \) satisfies
\[
p_y(x, y) L = -1
\]

The losses suffered by the man and the woman, \( ML \) and \( FL \), respectively, are given by
\[
ML = x + p(x, y) fL
\]
\[
FL = y + p(x, y)(1 - f) L
\]

where \( f \) is the fraction of the cost of the child borne by the man, \( 0 \leq f \leq 1 \).

**Noncooperative Decision Making**

The couple could make decisions about contraceptive care noncooperatively or cooperatively, that is, independently or jointly. Although cooperative decision making may be the more common case, it is useful to consider the noncooperative case as well, because transactions costs and strategic behavior mean that couples may not always act cooperatively. The noncooperative case will also serve as a reference point from which to compare the cooperative case. In considering the noncooperative case, we can apply what Cooter (1982, 18) calls the “normative Hobbes theorem,” which is the view that “the role of law is to minimize the inefficiency that results when bargaining fails” or, put differently, to maximize the value of the noncooperative solution of the game being played.
In the noncooperative game, the man’s welfare is, excepting corner solutions, maximized by setting $x$ so that it satisfies

$$p_x(x, y) = -1/f$$

while the woman sets her level of contraceptive care, $y$, so that, again excepting corner solutions, it satisfies,

$$p_y(x, y) = -1/(1 - f).$$

We can now analyze the effect of child-care laws (or “rules”) on contraceptive care.

One possible rule gives the woman the obligation to pay all the costs of child care. Call it the maternal-liability rule. If we wanted, we could also describe it as a rule that gives the father—and the child, the government, and everyone else—the right to have the woman look after the child. Mathematically, it is represented by setting $f = 0$. In this case, the man’s welfare is maximized by taking no contraceptive care (setting $x = 0$, a corner solution to the minimization problem). The woman then sets $y^*$, the optimal level for her under the maternal-liability rule, so that it satisfies

$$p_y(0, y^*) = -1.$$ 

Noncooperative decision making under this rule generates an inefficient result. The man takes too little care from the perspective of the couple (and society). Similarly, the woman’s level of care, while optimal given the man’s level of care, is now too high in the sense that the same probability of contraception could have been obtained at a lower cost. Formally, we know that $p_y(0, y^*) = p_y(x^*, y^*)$, since they are both set equal to the same value, namely $-1/L$. Since by assumption $x^* > 0$, $p_{yx} > 0$, and $p_{yy} > 0$, it follows that $y^* < y^m$.

An alternative rule would be to give the man the obligation to pay for the costs of the child or, equivalently, to give the woman—and everyone else—the right to have the man look after the child. Call this the paternal-liability rule. Mathematically, it is represented by $f = 1$. Analogous problems afflict this rule. The woman sets her level of care too low, and the man compensates by taking more care.

The problem here is that care should, in general, be taken by both parties. Rules that place liability on one party alone, such as maternal and paternal liability, cannot induce both parties to take optimal care. They are the equivalent of the rules of strict liability and no liability in tort law. Indeed, we could have described them as such had we taken the perspective of one of the parties. From the woman’s perspective, for example, maternal liability is strict liability and paternal liability is no liability. Our result, then, is an
application of Shavell's demonstration that the strict-liability and no-liability rules are suboptimal when both the injurer and the victim should take some care to avoid accidents.

A rule that divided the costs of child care between the man and the woman might appear a reasonable alternative to the extremes of no liability and strict liability. And it is this rule that most closely represents a government policy of requiring an absent parent to make child-support payments to the caregiving parent. Call this rule a joint-liability rule. It is the analogue of Shavell's strict-division-of-costs rule. Mathematically, it is represented by \( 0 < f < 1 \).

Unfortunately, this rule also fails to generate the optimal levels of contraceptive care. For \( 0 < f < 1 \), the value of \( x \) that satisfies \( p_x(x,y)L = -1/f \) will be less than the value that satisfies \( p_x(x,y)L = -1 \), since \( p_x(x,y) > 0 \).

The same argument applies to contraceptive care by the woman. Thus under the joint-liability rule, both the man and woman take less contraceptive care than is optimal. The probability of childbirth is therefore too high.

Are there any legal rules that would generate efficient contraception in a noncooperative game? The preceding analysis suggests that the following rule would work: make one of the parents strictly liable for the costs of child care and if a child is borne impose a tax equal to its cost on the other. Call this the tax-plus-liability rule. Suppose the woman is liable and the man pays the tax, \( T \). Then

\[
ML = x + p(x,y)T, \quad \text{and} \\
FL = y + p(x,y)L
\]

The man sets \( x' \), the optimal value of \( x \) under this rule, so that it satisfies

\[
p_x(x',y)T = -1, \quad \text{and}
\]

the woman sets \( y' \) so that it satisfies

\[
p_y(x,y)L = -1
\]

The woman's problem is the same as the problem we solved to find \( y^* \), so \( y' = y^* \). Since \( T \) is set to equal \( L \), the man's problem is the same as finding \( x^* \), so \( x' = x^* \) as well. Intuitively, the tax-plus-liability rule works because both the man and the woman take account of the full cost of children in making their contraceptive decisions.

Shavell shows that there are also several efficient nontax rules in the case of standard accidents, including a negligence rule, strict liability with the defense of relative negligence, negligence with the defense of contributory negligence, and comparative negligence.
All of them work by inducing one of the parties to take optimal care; the other then has complete liability for any losses and therefore also takes optimal care. All of these rules, however, require a court to determine whether at least one of the parties in an accident took at least the optimal level of care. While these rules might conceivably be used to improve contraceptive decisions, we will not consider them. They appear difficult to use in practice because of the difficulty of ascertaining the relevant facts, such as what contraception was in fact used and what would have been optimal. Accidental conceptions typically have few disinterested witnesses.

Before considering cooperative decision making, we can consider what the effects of the different rules are in the case that optimal contraceptive decision involves either the man or the woman taking all the care, rather than each taking some. This would be a natural solution if we were modeling contraceptive care as a dichotomous variable—use contraception or do not use contraception—rather than as a continuous variable, as we have been. In this dichotomous case, the man should use contraception if male contraception is more effective than female contraception (if $p(1,0) < p(0,1)$) and the woman should use contraception if female contraception is more effective than male contraception (if $p(1,0) > p(0,1)$).

Now either the maternal- or the paternal-liability rule is efficient. If male contraception is more cost-effective than female, a paternal-liability rule is efficient. If female contraception is more cost-effective than male, a maternal-liability rule is efficient.

This is a special case of a general principle that applies to many areas of law. If transaction costs or strategic behavior inhibit cooperation, the efficient legal rule is the one that assigns a right to the person who values it most. This formulation of the principle perhaps applies most obviously to property law, but it also applies to tort law. Define the right in the problem at hand as the right to have the other party look after the child. If male contraception is more cost-effective than female contraception, the right is more valuable to the woman than to the man: the right saves her more in contraceptive costs than it saves the man. Thus the efficient solution gives the woman the right or, in other words, makes the man liable for the costs of the child. (An alternative formulation of the principle that applies more naturally to tort law is this: assign liability to the least-cost avoider of risk, or in the case at hand the least-cost
contraceptor.) Let's turn now to the case of cooperative decision making.

**COOPERATIVE DECISION MAKING**

Suppose first that there are no transaction costs affecting cooperation between the man and the woman and that strategic behavior does not prevent cooperative solutions. What are the effects of the different rules we have considered: maternal liability, paternal liability, joint liability, and tax plus liability? We consider the simple liability laws first and then the hybrid tax-plus-liability rule.

Before Coase, it might have been thought that choice between different liability rules that distributed the cost of the accident in different ways would almost certainly influence contraceptive decisions—that paternal liability would encourage men to use contraception, for example. As Coase and his subsequent interpreters have shown, though, the mother and father will in the absence of transaction costs take the optimal amount of contraceptive care under the maternal-, paternal- and joint-liability rules. When they decide cooperatively and in the absence of transaction costs, they minimize $CL$, their joint loss function, and thus choose $x^*$ and $y^*$ under any of these three liability rules.

The normative Coase theorem can now be contrasted with the normative Hobbes theorem. Unlike the normative Hobbes theorem, the normative Coase theorem assumes cooperative decision-making and, in particular, that bargaining parties will not forego potential surpluses from trade by acting strategically. It is the view that we should "structure the law to remove the impediments to private agreements" (Cooter and Ulen 1988, 101)—in particular, we should minimize the costs of the transactions that need to take place to reallocate rights to the parties that value them most. We can see how the "theorem" might apply by imagining rules that would violate it. Suppose, for example, that neither potential parent was liable. In that case, either the potential child would be liable or someone else such as the government would be liable. Transaction costs between the unborn child and the parents are effectively infinite; so a child-liability law would not satisfy the normative Coase theorem. Transaction costs between the government and the potential parents are lower, but they are still probably higher than between the potential parents. So this liability rule probably fails the normative Coase test, too. Under most circumstances, we
would expect the paternal-, maternal-, and joint-liability laws to do well when judged by the normative Coase test, since the parents can transact with each other relatively easily. (Later, we consider the implications of assuming that the cost of transactions between the potential parents are important, though still low enough to make cooperative decision making a possibility.)

While paternal, maternal, and joint liability are all efficient, the choice between them may affect the distribution of welfare. A rule of paternal liability, for example, gives the mother a stronger starting position, or threat point, in bargaining over contraceptive care; but so long as they choose cooperatively they will choose optimally. The man will pay the woman to take the level of care that is best for both of them jointly (assuming again that $x^*, y^* > 0$). Under maternal liability, she will have to pay him something.

If we broaden our perspective of the problem at hand, however, it is less clear what will be the distributional consequences of different liability rules. In a more complicated model, we would need to consider the man and woman's decisions about whether and with whom to have sex. In the presence of a maternal-liability rule, for example, women will be less eager to have sex. Men may therefore have to give up part or all of the distributional gains we have just considered in order to persuade women to have sex. In general, the effect of legal rules on the distribution of income is diminished when the parties are in a contractual or market relationship with each other (Polinsky 1989, 121–127).

This argument, even if true in a traditional economic model of contractual decision making, is likely to strike some people as unconvincing. It might well be argued that women would not make up in contracting over sex what they lost in a change from paternal to maternal liability. Such arguments are hard to address in a standard economic analysis, and we will not pursue the subject in any detail. Sen (1990), however, sketches the outlines of an analysis of intrahousehold bargaining between men and women that may provide a foundation for the intuition that liability rules affect the distribution of welfare in a contractual setting. Adapting Sen's analysis to the problem at hand, we first assume that a person's well-being may differ from a standard preference-based measure of her utility. Then we note that liability rules may affect perceptions of fairness: under a long-standing maternal-liability rule, for instance, it may be believed that women have a moral duty to pay
the costs of children. If women believe they have such a moral duty, they may change their valuation of paying the costs of children; in terms of their preferences, the costs may be offset by the benefit of doing the right thing. If the moral benefits perfectly offset the costs, the maternal-liability rule does not alter women's approach to bargaining over sex.

In any case, let us suppose for the moment that the different liability rules do have different distributional consequences. If \( L \), the loss represented by child care, is unrelated to the distribution of wealth between the mother and the father, the level of contraceptive care and thus the probability of childbirth is invariant to the liability rule. If, on the other hand, \( L \) changes with the mother and father's wealth, the distributional effects might affect the probability of childbirth. However, the probability remains pareto-efficient.

Thus cooperative decision making solves the efficiency problems associated with maternal, paternal, and joint liability. It also destroys the efficiency of the tax-plus-liability rule. Under this rule, assuming that the tax is set equal to \( L \), the couple's joint loss is

\[
CL = x + y + p(x,y)2L
\]

The couple will therefore choose \( x_t \) and \( y_t \) so that they respectively satisfy

\[
p_x(x_t,y)L = 1/2, \quad p_y(x,y)L = 1/2.
\]

They now take contraceptive care up till the point where its marginal benefit is equal to half its marginal cost: they take too much contraceptive care, that is, and the probability of childbirth is inefficiently low. Why does the tax rule work in the noncooperative but not the cooperative case? In the former, it works to internalize an external cost. Cooperation also works to internalize all costs, however, and thus a tax in a cooperative setting is detrimentally distorting.

What happens when the transactions between the potential parents are costly? If they are high enough, cooperative decision making will not be efficient, and we return to the case of noncooperative decision making discussed earlier. Specifically, cooperative decision making will occur only if the transactions costs are less than \( CL(x^*,y^*) = CL(x^*, y^*) \), where the superscript \( n \) refers to the noncooperative levels of care given the liability rule in force. Suppose, however, that transaction costs are positive but not necessarily more than \( CL(x^*,y^*) = CL(x^*, y^*) \) and that they vary
according to which party initiates the bargaining that leads to a cooperative solution. Calabresi and Melamed (1972), citing an earlier work by Calabresi, *The Costs of Accidents* (1970), suggest an interesting rule for choosing among liability laws in this case. Calabresi and Melamed frame the rule as follows:

In particular contexts like accidents or pollution this suggests putting the costs on the party or activity which can most cheaply avoid them; . . . in the absence of certainty as to who that party or activity is, the costs should be put on the party or activity which can with the lowest transaction costs act in the market to correct an error in entitlements by inducing the party who can avoid social costs most cheaply to do so. (1096–1097)

To consider their rule, suppose also that the optimal contraceptive solution involves one party taking all the care, but that the government cannot tell which party that is, so that it cannot employ the rule of assigning liability to the least-cost avoider of risk. To make this plausible, imagine, first, that the cost-effectiveness of contraception varies among individuals, so that in some couples the man should take all the contraceptive care and in others the woman should take all the contraceptive care; and, second, that the government cannot observe the characteristics of individuals that determine the cost-effectiveness of their taking contraceptive care.

Does this rule have any application to the problem of accidental children? It would if social custom made it easier for either the man or the woman to initiate bargaining. For concreteness, suppose that women are discouraged by social custom from questioning the status quo. As a result, let us assume, the transaction costs of female-initiated bargaining are always greater than $C_l(x^*, y^*) - C_l(x^*, y^n)$ while the transaction costs of male-initiated bargaining are less than $C_l(x^*, y^*) - C_l(x^n, y^n)$ when and only when the liability rule is inefficient for the couple under consideration.

In this case, it seems that Calabresi and Melamed would choose paternal liability. But, if we have interpreted their rule correctly, it is not clear that it matters. To see the problem with the rule, consider paternal and maternal liability in turn. Under paternal liability, if the man is the least-cost contraceptor, the efficient outcome occurs without bargaining: the man should use contraception and the rule induces him to do so. If the woman is the least-cost contraceptor, there are potential gains from trade, and the man will pay the woman to take on the contraceptive burden. Either way, the right
party uses contraception. Now consider the maternal-liability rule. If the woman is the least-cost contraceptor, she should use contraception and the rule induces her to do so. If the man is the least-cost contraceptor, the man should use contraception; but the woman will not initiate bargaining, because the transaction costs would be, by assumption, greater than the gains from trade. Perhaps this is why Calabresi and Melamed suggest their rule for choosing between legal rules. But the man would initiate bargaining in this case and offer to use the contraception if the woman made him a payment of an amount less than her cost of using contraception but greater than his. Thus the choice of rule is unimportant, and Calabresi and Melamed's rule appears to add no value.

CONCLUSIONS
What conclusions can we draw from this brief analysis of the effects of child-care laws? Most important to the analysis is whether contraceptive decisions are made cooperatively or noncooperatively.

If they are made noncooperatively, no straightforward liability rule is efficient in general. Complete paternal liability will lead the woman to take no care. Complete maternal liability will lead the man to take no care. The more common rule of joint liability leads both the man and the woman to take too little care. The only rules that in general produce efficient outcomes are negligence rules, which are likely to be impractical in the case at hand, and a rule that in effect makes both the man and the woman liable for all the costs of children, for example, by making one liable and taxing the other an equivalent amount.

Under cooperative decision making, however, all the rules are efficient, except for the tax-plus-liability rule, which leads to excessive contraceptive care.

Is it cooperative or noncooperative decision making that is the more realistic assumption, then? The question is largely beyond the scope of this note, but we can make a few observations. The cases in which noncooperative decision making is usually considered likely are those involving many parties (a firm considering whether to emit pollution over a city) and those involving strangers (a driver considering how much care to take to avoid hitting a pedestrian). In other cases, cooperative joint-profit-maximizing decisions seem likely. Since contraceptive decision making involves two people only—two people who are not complete strangers—cooperative
decision making may be expected. Indeed, Hoffman and Spitzer (1982) find in an experimental setting similar to the choice of contraceptive care under different liability rules that individuals act cooperatively under both full and incomplete information. On the other hand, contraceptive bargaining may sometimes be subject to transaction costs that do not beset other cases. For example, a recently acquainted couple may be embarrassed to talk about sex and contraception. In addition, explicit bargaining over each party's obligations tends to imply a lack of mutual trust. In related contexts, some researchers have argued that treating individuals as untrustworthy actually makes them less reliable and honest than they would have been had they been treated in a trusting manner. As a result, it is said, it may be optimal for principals not to intensively monitor agents with whom they have a close personal relationship. If what might be true of monitoring were true of contracting, explicit bargaining over contraception would have extra costs; but at the same time trust might sometimes facilitate the realization of the cooperative solution without explicit bargaining. Low-cost bargaining also requires that the parties can easily make side payments to each other, so that the realization of a cooperative surplus does not make one of the parties worse off than in the absence of cooperation. In the case of agreements between close friends, explicit side payments tend to be socially disfavored (that is, are more costly); other less-explicit forms of redistribution, such as payments in kind or gift giving, may be more socially acceptable, but in contrast to the pecuniary side payments of traditional game theory, they are not generally costless. It is not obvious, therefore, how often cooperative decision making about contraception is likely to occur.

Notes
1 See Cooter (1986), for example, for an explication of the general point.
2 See Williamson (1993), for example.

References


ONE MAN, ONE GRANT: EQUITY IN FEDERAL AID TO STATES

Karen Baehler*

Policy makers in Washington are moving swiftly to change the way federal funds flow to states. This paper describes several diagnostic tools that analysts can use to evaluate the distributional fairness of both existing and proposed approaches to federal grants-in-aid. When the tools are applied to four small grant programs in the Department of Health and Human Services, the results suggest that capped programs are more equitable than open-ended programs. The author argues for wider use of these diagnostic techniques to evaluate the equity of various federal funding vehicles.

With President Clinton's promise to "end welfare as we know it" and the commitment of the Republican Congress to end federal spending as we know it, the time is ripe for major transformations in the federal-state partnership to aid the poor.

Many of the early reform proposals from the 104th Congress have focused on changing the way in which social-welfare funds flow from the federal government to the states. The "unfunded mandates" bill, passed in January, requires that the costs of certain new federal requirements for states be subsidized with federal aid. Other proposals in the same spirit include replacing as many as 350 existing categorical federal social programs—for everything from school lunches to crime fighting and public transportation—with lump-sum block grants. Supporters of these ideas generally argue for them on grounds of economy and efficiency and claim that block grants will save money and improve services by reducing

*M.P.P., Ph.D. Candidate, School of Public Affairs, University of Maryland; B.A., St. Olaf College.
program duplication, cutting federal red tape and devolving control to state decision makers who are more likely to experiment and innovate.

The debate to this point has largely overlooked equity: Will the new block grants be distributed fairly among the states? Will the new burdens imposed by cuts in federal reimbursements be distributed fairly? At first glance, these questions may seem premature. Common sense dictates that Congress should decide first how much to spend, then design the mechanism for funding, and only subsequently fight about the details of distributing the funds to the states. As this paper seeks to show, however, decisions about spending and funding mechanisms determine, to a great extent, the final distribution of resources. In other words, issues of equity are embedded in arguments over economy and efficiency; they should not be postponed until later in the debate.

Politicians know this, of course. State officials have a rough idea of which federal funding arrangements provide their states with the best cash flow. They decide whether to support or oppose reform proposals based on these calculations. Many of the most intense battles in Congress are between those states and regions of the country that expect to gain from change and those that expect to lose. Policy analysts, however, do not have clear methods for evaluating equity. We can predict the effects of a reform on the distribution of resources to states and identify winners and losers. But we have a hard time determining whether the new distribution is more or less fair than the old one.

This paper begins to explore the equity implications of different types of federal funding, using the experience of four small social-welfare programs administered by the U.S. Department of Health and Human Services. The analysis points to some preliminary conclusions and suggests areas for further study. The preliminary conclusions can be summarized as follows:

- Federal aid to states is consistent with other components of the U.S. social-welfare system: it claims to seek greater equality in the distribution of resources, but in fact contributes very little to equality.
- The main reason for this disjunction between goal and outcome is the conflict between the distributive principles embedded in particular funding formulae, specifically the principles of need and contribution.
Capped spending programs and programs that do not require state matching funds appear to come closer to the goal of egalitarian distribution than open-ended spending programs and programs with matching requirements.

DEFINING EQUITY

The first step in determining the fairness of a particular distribution of resources is to define standards of equity. Philosophical theories of justice provide benchmarks against which to measure the equity of a particular distribution. Philosophers, however, disagree about the meaning of fairness and offer the policy analyst few tools for choosing among their theories.

Another place to look for clues about the meaning of equity is the program to be analyzed. Programs often contain their own sets of principles of fair distribution. The programs themselves and the principles they embody derive legitimacy from the fact that they are the product of political deliberation and consensus. The principles that emerge can be used to evaluate the program’s outcomes: Does the program honor its own internal principles? Does it achieve its stated goals? This approach to equity, known as local justice, acknowledges that different programs with different goals may call for different standards of fairness to be applied (Young 1994; Elster 1992).

Michael Walzer provides philosophical grounds for this pragmatic approach. His theory of “complex equality” is based on the notion that standards of fair distribution are intrinsic to the social meanings of the goods being distributed. For any particular good, Walzer argues, “If we understand what it is, what it means to those for whom it is a good, we understand how, by whom, and for what reasons it ought to be distributed. All distributions are just or unjust relative to the social meanings of the goods at stake” (Walzer 1983, 9).

In the case of grants-in-aid to states, standards of equity or fairness can be found on at least two levels: the level of general justification for federal aid to states and the level of specific program structure and formulae.

Fundamentals

Consider first the category of grants-in-aid in general. What are their social meanings? According to Alice Rivlin, current director of the U.S. Office of Management and the Budget and former director of the Congressional Budget Office, these programs were estab-
lished to serve two basic functions: reducing inequalities among states and exerting quality control over state programs. In Rivlin’s words, “The principle of using federal spending programs to accomplish redistribution across state lines is an accepted feature of American federalism. Usually these programs have combined the two objectives of improving services and equalizing them” (Rivlin 1992, 136).

Applied to the area of social welfare, the idea seems to be that state residency should not confer advantage or disadvantage on those seeking social services. Services of equal quality should be equally accessible in every state. This is as close to an explicit principle of egalitarian distribution as you will find in U.S. social policy.

**Specifics**

Turning from general justification to the details of specific social-services funding programs reveals important limitations to the version of egalitarianism in operation. This article focuses on the funding mechanisms for four federal social-services grant programs targeting the needs of children and their families. The programs were selected to represent at least a few of the primary funding arrangements used by the federal government to assist states in the area of social services to the poor. Better-known poverty programs like Aid to Families with Dependent Children (AFDC), Supplemental Security Income (SSI), and food stamps were not included because they are based on direct transfers of cash or cash-like vouchers to individual recipients rather than to state agencies functioning as service providers.

The analytical techniques demonstrated here can and should be applied to dozens of other federal social-services grant programs. Examples of other candidates for equity analysis include the Social Services Block Grant, the school lunch program, compensatory education grants, various housing grants, Head Start, and many other programs. The following programs are analyzed:

1) **Child Welfare B**: This label refers to Title IV-B of the Social Security Act which established the Child Welfare Services Program, some version of which has been in operation since 1935. IV-B provides federal grants to help states pay for a wide variety of services to address problems related to neglect, abuse, exploitation, or delinquency of children.
2) *Child Welfare E:* This label refers to Title IV-E of the Social Security Act, which established the AFDC Foster Care Program. IV-E reimburses states for the costs of keeping neglected and abused children who are eligible for AFDC in foster care. The federal government has been providing such reimbursements since 1961.

3) *Child Care Block Grant:* This label refers to the Child Care and Development Block Grant created in 1990. This block grant helps pay for child care services for low-income families and funds state activities to improve the supply and quality of child care in general.

4) *Child Care AFDC:* This label combines two programs, AFDC child care and transitional child care, under which the federal government reimburses states for the costs of child care for: (a) AFDC recipients who are either working or participating in a state-approved education and training program; and (b) former AFDC recipients who have worked themselves off the welfare rolls (available up to 12 months after the AFDC benefits stop).

These four programs differ in both scale and pattern of distribution. As Table 1 shows, Child Welfare E is relatively large, with annual expenditures more than six times that of the smallest program, Child Welfare B.

<table>
<thead>
<tr>
<th>Table 1: FY 1991 Federal Appropriations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child Welfare E</strong></td>
</tr>
<tr>
<td><strong>Child Care Block</strong></td>
</tr>
<tr>
<td><strong>Child Care AFDC</strong></td>
</tr>
<tr>
<td><strong>Child Welfare B</strong></td>
</tr>
</tbody>
</table>

Table 2 presents a rough outline of the funding mechanisms employed by each of the four programs to distribute funds to states.
Table 2: A Sample of Federal Funding Mechanisms for Social Services Grants to States

<table>
<thead>
<tr>
<th>Program</th>
<th>Budgetary Classification</th>
<th>Allocation Formula</th>
<th>Target Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Welfare B</td>
<td>authorized and appropriated by Congress</td>
<td>• state grant based govt. doesn't require* on state's population &lt; age 21 and per capita income • state must provide 25 percent matchfed.</td>
<td></td>
</tr>
<tr>
<td>Child Welfare E</td>
<td>open-ended entitlement to states</td>
<td>state reimbursed at Medicaid rate **children eligible for AFDC</td>
<td></td>
</tr>
<tr>
<td>Child Care Block Grant</td>
<td>authorized and appropriated by Congress</td>
<td>state grant based on state's: • proportion of children &lt; age five • no. of children getting free lunch subsidy • per capita income; no state match required</td>
<td></td>
</tr>
<tr>
<td>Child Care AFDC</td>
<td>open-ended entitlement to states</td>
<td>state reimbursed at Medicaid rate** certain current and former AFDC recipients</td>
<td></td>
</tr>
</tbody>
</table>

*Although the federal government does not impose eligibility requirements for these programs, states use the funds mostly for services to low-income families and individuals.

**Each state's federal Medicaid matching rate is calculated based on the state's per capita income, with a floor of 50 percent. The current range of state rates is between 20 and 50 percent. The average rate is 46 percent. The Medicaid allocation rate can be described either as a state matching rate (average of 46 percent) or a federal reimbursement rate (average of 54 percent).

The kind of egalitarianism represented by federal grants-in-aid for social services now comes into sharper focus. It is basically a population-weighted egalitarianism based on the well-known normative principles of need and contribution.
Distribution based on need

Only one element can be found in all four formulae: per capita income. Under the open-ended entitlement programs, Child Welfare E and Child Care AFDC, states are reimbursed for their expenditures at the Medicaid rate, which is based on per capita income. The two nonentitlement programs, Child Welfare B and Child Care Block Grant, include state per capita income as one of several components in their distribution formulae.

Per capita income reflects the dual needs faced by poor states. First are the needs of potential beneficiaries of social-welfare programs. States with proportionally more low-income people are expected to face greater demand for public social services. Second are the needs of taxpayers. Low-income states, in the absence of other revenue opportunities, have less ability to raise money through income taxes to pay for public social services. For both of these reasons, low per capita income states seem to have a stronger claim on federal aid than high per capita income states.

In addition to per capita income, the nonentitlement programs include other measures of potential demand for services, including each state’s child population and child poverty population (via subsidized lunches). All of these arrangements clearly are intended to funnel more dollars per person to poorer states. The dominance of per capita income in Table 2 lends support to Rivlin’s argument that federal aid is intended to equalize state resources.

Distribution based on contribution

Three of the four programs, excluding the Child Care Block Grant, require states to contribute some portion of the costs of services. The Child Welfare B program has a flat matching rate of 25 percent. The two open-ended entitlement programs, Child Welfare E and Child Care AFDC, use the progressive Medicaid matching rate which requires a smaller state contribution per federal dollar for poorer states. Under these entitlement programs, states are reimbursed for a portion of all eligible expenses, with no limits. Matching requirements are intended to prevent states from substituting federal dollars for their own spending without increasing the quantity or quality of services.
Conflicts

The information in Table 2 allows the extraction of principles of distribution that are relevant here, but it does not contribute to an understanding of how the principles will interact. The disciplines of philosophy and public policy analysis suggest two areas of likely conflict.

Kenneth Boulding describes the first problem, a clash between principles.

The concept of justice has two incompatible dimensions, and one suspects that the dynamics of any prevailing ideology of justice is a result of a constant dialectical interaction between the two. One is the dimension of distribution according to desert, in which we have injustice if one person gets more than he deserves and another gets less. The other dimension is that of distribution according to need. This is the principle of commensality, that we are all sitting around the great table of the universe and the goodies should be dished out to everybody irrespective of desert. We are all equal in the sight of God, or whatever is the current ideological substitute for Divinity. Then we run into the dilemma that if people get what they deserve, some people would not get what they need, and if people get what they need, some people would not get what they deserve. (Boulding 1976, 18)

Boulding predicts that a distribution function based on need might yield very different allocations than a distribution function based on contribution. Indeed, to the extent that need is inversely proportional to contribution, the two final distributions would be radically different. In the case of systems that employ both principles, as in the funding formulae analyzed for this article, the features of the final distribution are unclear. Which principle will prevail in these distributions? The analysis reported later suggests that, under conditions of open-ended funding, the principle of contribution tends to eclipse the principle of need.

A recent analysis of the Maternal and Child Health Block Grant formula illustrates the second problem, a clash within the principle of need between the needs of social-service beneficiaries and the needs of state taxpayers (U.S. GAO 1992). That study found that "equity for state taxpayers and equity for children cannot be fully satisfied simultaneously." The analysts constructed two separate allocations of program funds, the first based on a measure of each state's population in need of maternal and child health services, the
second based on a measure of each state's ability to raise revenue to pay for these services. They found significant divergence between the two allocations. According to the study, the only way to reconcile the conflict is to employ a composite formula that strikes a compromise between the two types of need.

The programs analyzed in this paper sidestep the conflict identified by GAO by using a single measure, per capita income, to capture both concepts of beneficiary need and taxpayer need. For a variety of reasons, however, per capita income may be a poor proxy for either definition of need. The assumption that per capita income correlates closely with poverty rates and tax capacity is flawed. Thus, the conflict between beneficiary equity and taxpayer equity may be lurking beneath the surface of these four programs, indicating the need for further research.

At the same time, the conflict appears to be complicated in three of the four programs by the operation of matching requirements, the contribution component of the formulae. Matching requirements, particularly in the case of open-ended reimbursement funding, favor states that can afford to spend more on clients up front. They reflect the principle that states that spend more of their own funds should get more federal aid. This element seems to contradict the idea of taxpayer equity—that states with less ability to tax, and therefore less ability to pay for services themselves, should receive more federal aid. If taxpayer need (inability to tax and pay) is the inverse of contribution (ability to pay up front), then we might expect these two elements of the federal funding formulae to cancel each other out, leaving beneficiary need as the dominant force driving distribution.

Conspicuous Absence

It is also worth noting a prominent principle of justice that does not appear in these formulae: the principle of efficiency. The fact that one state is likely to achieve more improvement in client well-being per federal dollar than another state is not considered to be valuable or appropriate information in this context.\(^5\)

There are good reasons for leaving efficiency out of the distribution function for grants-in-aid. The most obvious is the technical problem of measuring and comparing efficiencies across states. There is also the problem of opening a new door to potential political manipulation and corruption. If decisions about federal
funding turned on the results of evaluations of state social-services agencies, a lot of time and energy would have to be spent safeguarding the integrity of the evaluation process. Third, there is historical precedent against such a calculation. Distributions of federal funds to states have rarely been made on the basis of agency efficiency or effectiveness, particularly in an area like social services that some consider to be part of the essential safety net. Finally, funding based on efficiency would penalize the residents of a state with reduced federal grants because that state's social-services agency is below par, an idea unlikely to gain political support.

The apparent lack of concern about efficiency in the funding formulae reflects, in part, the difficulties in measuring the outputs of a social-welfare system. Without any widely accepted, comprehensive measure of child welfare, it is difficult to determine whether children in Wisconsin, for example, are better or worse off than children in Michigan. Likewise, it is difficult to determine whether or not the social-services system in Wisconsin is producing greater improvements in child-welfare per federal dollar than its counterpart in Michigan. Unfortunately, child-welfare experts are a long way from having such a comprehensive measure. Analysis of the outputs of the state social-welfare systems is far beyond the scope of this paper. The following discussion looks only at equity in federal funding inputs to these systems, recognizing that an equitable distribution of inputs does not guarantee an equitable distribution of outputs.

**Measuring Equity**

The picture that emerges from this examination of the federal social-services distribution function is of a pair of potentially conflicting principles of justice set within a roughly egalitarian framework. What kinds of actual distributive outcomes can we expect from this system? Which principle will prevail? How equitable will the final allocations be?

The simplest and clearest way to pursue these questions is to separate the equity principles into three parts—beneficiary equity, taxpayer equity, and contribution—and evaluate program distributions according to each standard. This section presents the results from the first step in that research project, based on a disaggregation of the standard of beneficiary equity. The analysis employs
three simple techniques to measure the extent to which the distributions produced under the four funding formulae described above reflect the principle of beneficiary need. The results provide insights into the relationships among principles of equity as they operate in the world of federal social-welfare funding.

**Need as Poverty**

In order to examine the role of beneficiary need in the distributions, we must first operationalize the concept by asking the question: Who are the potential beneficiaries for the four programs, and what do they have in common? The answer is: low-income children and their families. Two of the four programs are restricted by law to clients who either are receiving, have received, or are eligible to receive AFDC. A third program is explicitly targeted to children from poor homes. The fourth program, although not formally means-tested, in practice provides most of its services to low-income families.\(^{10}\)

Thus, the simplest measure of beneficiary need that we can apply to all four programs seems to be child poverty. The child poverty count offers a more direct gauge of a state's at-risk pool than does per capita income, which reflects the entire distribution of income and does not accurately predict the size of the bottom group. The child poverty count also avoids the problem of trying to measure two things at once, like beneficiary need and taxpayer need in the case of per capita income.

Of course, the poverty line is an imperfect standard of beneficiary need because it does not take into account the distribution of income among the poor, and it defines poverty narrowly in terms of income only.\(^ {11}\) Perhaps most problematic for our analysis, the U.S. poverty line also makes no allowances for differences in the cost of living across the states. Thus, it may inflate the poverty head count in states where the cost of living is low. Despite these problems, the poverty line is widely accepted and widely used in empirical studies. It is the best choice for this study because it provides a uniform national standard for comparison. In further analyses, the author plans to test the sensitivity of the results obtained here to the choice of denominator.\(^ {12}\)

**Dollars Per Needy Child**

One crude measure of equity proportional to the population at risk of needing services is that of federal spending per poor child.
For this calculation, we simply divide each state's total federal allocation for a given program by that state's total population of poor children.

The data for this exercise come from two sources. The numerators, federal allocations to states for various programs, are reported in the *Green Book*, an annual publication of the U.S. Government Printing Office that provides an overview of programs under the jurisdiction of the House Ways and Means Committee. The denominators, the number of related children under 18 years of age living in households with income below the poverty line, by state, are from the 1990 Census, *U.S. Summary of Social and Economic Characteristics*.

Figures for federal spending per poor child, for each program and each state, are summarized below.

Table 3. Federal Spending per Poor Child
Summary of FY 1991 Distributions (in dollars)

<table>
<thead>
<tr>
<th>Program</th>
<th>Child Welfare B</th>
<th>Child Welfare E</th>
<th>Child Care Block</th>
<th>Child Care AFDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>16.84</td>
<td>8.78</td>
<td>50.26</td>
<td>2.33</td>
</tr>
<tr>
<td>MS</td>
<td>23.81</td>
<td>157.89</td>
<td>61.26</td>
<td>28.71</td>
</tr>
<tr>
<td>High</td>
<td>44.41</td>
<td>614.38</td>
<td>95.42</td>
<td>145.40</td>
</tr>
</tbody>
</table>

* This is total U.S. spending for program X divided by the U.S. total of related children under 18 with incomes below poverty (roughly 11 million in the 1990 census).

Table 3 presents a situation that may be described as "transparent inequality" (Young 1994, 147). Although the distributions are the result of rational formulae, anyone can see that the outcomes heavily favor certain states over others. The differences between the low and high figures for each program are most dramatic, but they are not simply the result of extreme outliers. Large differences between the first and fourth quartile means for each program support the conclusion that each of these programs is a transparently unequal distributor of funds. In other words, for each of the four programs studied, poor children in some states receive vastly
more federal social-services funding than poor children in other states. I will call this the residency penalty.

The measure of federal dollars per poor child allows crude comparisons of inequality across programs. For example, Table 3 suggests that, among the programs analyzed, Child Welfare E (AFDC Foster Care) is the most unequal and the Child Care Block Grant is the most equal, based on the relative sizes of the gaps between the high and low figures. This judgment is based on the observation that the high figure for the Child Welfare E program (New York) is more than 60 times that of the low figure (Mississippi), and more than four times the mean. In contrast, the high figure for the Child Care Block Grant program (New Hampshire) is less than twice that of the low figure (Michigan), and about 150 percent of the mean.

The differences in scale of funding from program to program (see Table 1), however, make interpretation of these figures difficult. In order to make real comparisons across programs, a standardized measure is needed.

**Coefficient of Equity**

One standardized measure of equity is the ratio between the funding that state X actually receives under program Y (actual allotment) and the amount of funding that state X should receive under conditions of equity based on beneficiary need (the a priori fair share). If these state ratios differ dramatically one from the other, the distribution may be said to suffer from what Peyton Young calls “transparent inequity”: “An allocation is transparently inequitable if for some two claimants I and J, the ratio between I’s allotment [what he actually receives] and I’s share [what he deserves] is greater than the ratio between J’s allotment and J’s share” (Young 1994, 147).

The problem of fair distribution to 50 states is just a larger version of the two-claimant problem described by Young. An ultimately fair solution to the 50-state problem will mean that every two states divide their allotted share of federal aid equitably between them (Young 1994, 15). Viewed this way, the distribution function is an aggregation of separate equity relationships among pairs of individual states.

The next step in the analysis is to develop a specific model that adapts Young’s definition to the problem at hand. The model
proposed here is a simple calculation of each state's ratio of its actual allotment to its a priori fair share for each program.\textsuperscript{13} The formula is as follows:

\[ E_x = \frac{A_{xy}}{C_x \cdot T_y} \]

where \( E_x \) is state x's ratio of actual allotment to fair share (which I am calling the coefficient of equity), \( A_{xy} \) is state x's 1991 allotment of federal funds for program y, \( C_x \) is the proportion of U.S. poor children living in state x according to the 1990 census, and \( T_y \) is total federal funding for program y in 1991. When \( E < 1 \), the state's actual allotment is less than its fair share based on its proportion of the nation's poor children. When \( E > 1 \), the state's actual allotment is more than its fair share based on its proportion of the nation's poor children.

An example will help show how the coefficient of equity works. Take Alabama, which in 1991 received $5,634,000 in Child Welfare B funds. In the 1990 census, Alabama was home to 2.2 percent of the nation's poor children. Total U.S. spending for Child Welfare B in 1991 was $265,823,000. Using these figures, Alabama's ratio of actual Child Welfare B funding to its fair share (the coefficient of equity) is:

\[
\frac{5,643,000}{(0.022)*(265,823,000)} = 0.94
\]

Thus, Alabama receives slightly less under the existing allocation formula than it would under a formula that considered only the state's proportion of the nation's poor children. Its actual allotment is 94 percent of its fair share.

The coefficients of equity for the other states in the child-welfare programs are shown in Table 4a, listed from lowest to highest:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D.C.</td>
<td>.71</td>
<td>Mississippi</td>
<td>.06</td>
</tr>
<tr>
<td>Louisiana</td>
<td>.71</td>
<td>Alabama</td>
<td>.13</td>
</tr>
<tr>
<td>Mississippi</td>
<td>.72</td>
<td>Idaho</td>
<td>.16</td>
</tr>
<tr>
<td>Texas</td>
<td>.79</td>
<td>Arkansas</td>
<td>.20</td>
</tr>
<tr>
<td>New Mexico</td>
<td>.80</td>
<td>North Carolina</td>
<td>.20</td>
</tr>
<tr>
<td>New York</td>
<td>.82</td>
<td>Indiana</td>
<td>.23</td>
</tr>
<tr>
<td>California</td>
<td>.83</td>
<td>Hawaii</td>
<td>.26</td>
</tr>
<tr>
<td>Arizona</td>
<td>.87</td>
<td>Texas</td>
<td>.30</td>
</tr>
<tr>
<td>State</td>
<td>Proportion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kentucky</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arkansas</td>
<td>.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alabama</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>West Virginia</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td>.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illinois</td>
<td>.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Dakota</td>
<td>1.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Montana</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Hampshire</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Carolina</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Massachusetts</td>
<td>1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connecticut</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>1.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Carolina</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alaska</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansas</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevada</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Dakota</td>
<td>1.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minnesota</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iowa</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maryland</td>
<td>1.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idaho</td>
<td>1.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>1.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhode Island</td>
<td>1.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebraska</td>
<td>1.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wyoming</td>
<td>1.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td>1.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaware</td>
<td>1.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hawaii</td>
<td>1.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utah</td>
<td>1.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vermont</td>
<td>1.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wyoming</td>
<td>.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Dakota</td>
<td>.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utah</td>
<td>.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Mexico</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Carolina</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arizona</td>
<td>.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorado</td>
<td>.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td>.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td>.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>West Virginia</td>
<td>.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td>.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevada</td>
<td>.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delaware</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhode Island</td>
<td>.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebraska</td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Dakota</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kentucky</td>
<td>.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illinois</td>
<td>.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alaska</td>
<td>.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iowa</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansas</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Montana</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.C.</td>
<td>1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Massachusetts</td>
<td>1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minnesota</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td>1.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td>1.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>1.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Hampshire</td>
<td>1.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>1.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>1.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connecticut</td>
<td>1.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vermont</td>
<td>2.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>3.89</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the goal of federal aid is to equalize the funds available for services to poor children across states, Table 4a provides clear evidence that Child Welfare B is the superior of the two programs. Under Child Welfare B, no state receives less than 71 percent of its proportionate share of federal dollars per poor child, and no state receives more than 186 percent.
In comparison, under Child Welfare E, half of the states receive less than half of their proportionate share of federal dollars per poor child, and one state receives almost four times its fair share. These calculations reveal that Child Welfare E operates, in practice, as a massive redistributor of federal dollars to the state of New York. In 1990, 7 percent of the nation’s poor children lived in New York, which entitles the state to 7 percent of federal Child Welfare E spending, or $123.5 million according to the definition of fairness being employed here. New York’s actual allotment of federal Child Welfare E spending in 1991 was $480.8 million or 27 percent of total expenditures. Vermont, too, received a relative windfall under the Child Welfare E program, but the scale of redistribution to Vermont, compared to New York, is much less. Vermont received roughly $4 million more than it should have, based on its share of U.S. poor children, compared to an over allocation of $350 million to New York.

Coefficients of equity follow similar, but less dramatic, patterns in the federal child-care funding programs, as shown in Table 4b. Here, as expected, the Child Care Block Grant is the superior vehicle for achieving equality in state funding. Under the AFDC Child Care program, Massachusetts received a windfall equivalent to about 5.5 percent of total program expenditures in 1991.

Table 4b. Coefficients of Equity in Child Care, FY 1991, in Increasing Order

<table>
<thead>
<tr>
<th>State</th>
<th>Child Care Block</th>
<th>State</th>
<th>Child Care AFDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan</td>
<td>.82</td>
<td>Mississippi</td>
<td>.08</td>
</tr>
<tr>
<td>D.C.</td>
<td>.83</td>
<td>South Carolina</td>
<td>.10</td>
</tr>
<tr>
<td>New York</td>
<td>.86</td>
<td>Missouri</td>
<td>.19</td>
</tr>
<tr>
<td>West Virginia</td>
<td>.88</td>
<td>Hawaii</td>
<td>.28</td>
</tr>
<tr>
<td>Ohio</td>
<td>.89</td>
<td>California</td>
<td>.29</td>
</tr>
<tr>
<td>Illinois</td>
<td>.91</td>
<td>Alabama</td>
<td>.39</td>
</tr>
<tr>
<td>California</td>
<td>.94</td>
<td>Arizona</td>
<td>.40</td>
</tr>
<tr>
<td>Louisiana</td>
<td>.94</td>
<td>Idaho</td>
<td>.55</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>.94</td>
<td>New Mexico</td>
<td>.59</td>
</tr>
<tr>
<td>New Mexico</td>
<td>.95</td>
<td>Illinois</td>
<td>.61</td>
</tr>
<tr>
<td>Texas</td>
<td>.95</td>
<td>Tennessee</td>
<td>.63</td>
</tr>
<tr>
<td>Arizona</td>
<td>.96</td>
<td>Texas</td>
<td>.63</td>
</tr>
<tr>
<td>Arkansas</td>
<td>.96</td>
<td>Ohio</td>
<td>.67</td>
</tr>
<tr>
<td>Montana</td>
<td>.96</td>
<td>West Virginia</td>
<td>.67</td>
</tr>
<tr>
<td>Kentucky</td>
<td>.97</td>
<td>Kentucky</td>
<td>.76</td>
</tr>
<tr>
<td>Mississippi</td>
<td>.97</td>
<td>Iowa</td>
<td>.78</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>.97</td>
<td>Virginia</td>
<td>.78</td>
</tr>
</tbody>
</table>
Tables 4a and 4b provide a basis on which to compare all four programs. In terms of equal allocations for states proportional to need, Child Welfare E and AFDC Child Care come out on the bottom. It is not entirely clear, however, which of the two is more unequal. The answer seems to depend on which end of the distribution is given the most weight. The AFDC Child Care program produces a large number of very big winners, most of which are small states. According to Table 4b, nine states receive more than double their fair share under this program; three states receive more than triple; and one state receives five times its fair share. Seven states receive less than half their fair share under AFDC Child Care. In contrast, the Child Welfare E program results in a large number of losers (25 states receive less than 50 percent of their fair

<table>
<thead>
<tr>
<th>State</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>.85</td>
</tr>
<tr>
<td>South Dakota</td>
<td>.87</td>
</tr>
<tr>
<td>Montana</td>
<td>.92</td>
</tr>
<tr>
<td>North Carolina</td>
<td>.95</td>
</tr>
<tr>
<td>Arkansas</td>
<td>.99</td>
</tr>
<tr>
<td>Colorado</td>
<td>1.01</td>
</tr>
<tr>
<td>Nevada</td>
<td>1.01</td>
</tr>
<tr>
<td>Michigan</td>
<td>1.14</td>
</tr>
<tr>
<td>Louisiana</td>
<td>1.18</td>
</tr>
<tr>
<td>Maine</td>
<td>1.18</td>
</tr>
<tr>
<td>Kansas</td>
<td>1.24</td>
</tr>
<tr>
<td>New York</td>
<td>1.30</td>
</tr>
<tr>
<td>Georgia</td>
<td>1.37</td>
</tr>
<tr>
<td>Florida</td>
<td>1.41</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1.56</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>1.58</td>
</tr>
<tr>
<td>Washington</td>
<td>1.70</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1.78</td>
</tr>
<tr>
<td>New Jersey</td>
<td>1.81</td>
</tr>
<tr>
<td>North Dakota</td>
<td>1.85</td>
</tr>
<tr>
<td>Oregon</td>
<td>2.04</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>2.11</td>
</tr>
<tr>
<td>Indiana</td>
<td>2.25</td>
</tr>
<tr>
<td>Connecticut</td>
<td>2.41</td>
</tr>
<tr>
<td>Delaware</td>
<td>2.43</td>
</tr>
<tr>
<td>Maryland</td>
<td>2.67</td>
</tr>
<tr>
<td>Minnesota</td>
<td>2.78</td>
</tr>
<tr>
<td>Utah</td>
<td>2.90</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>2.94</td>
</tr>
<tr>
<td>Nebraska</td>
<td>3.14</td>
</tr>
<tr>
<td>Vermont</td>
<td>3.50</td>
</tr>
<tr>
<td>D.C.</td>
<td>3.51</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>5.06</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Missouri 1.00
Colorado 1.04
Tennessee 1.04
Florida 1.05
Oklahoma 1.05
Wisconsin 1.05
Georgia 1.07
New Jersey 1.07
Rhode Island 1.08
Washington 1.08
Alabama 1.09
Nevada 1.09
Oregon 1.09
Connect 1.12
Virginia 1.12
Indiana 1.13
South Dakota 1.14
North Dakota 1.17
South Carolina 1.18
Iowa 1.19
Delaware 1.20
Minnesota 1.20
Alaska 1.21
Maryland 1.21
Wyoming 1.21
Maine 1.22
Kansas 1.23
Idaho 1.24
Nebraska 1.25
North Carolina 1.26
Vermont 1.27
Hawaii 1.41
Utah 1.43
New Hampshire 1.56
share) and two big winners (with two and a half times and almost quadruple their fair shares). The scale of the redistributions in the two programs points to Child Welfare E as the more egregious violator of equity. As noted above, Child Welfare E shifts an additional 20 percent of total program expenditures from the states at the bottom of the distribution to New York (on top of New York's fair share of 7 percent).

The two remaining programs appear to operate roughly within the bounds of what might be considered equitable distributions. Although neither program produces a strictly equal distribution of funds (in which, for every state, \( E_x = 1 \)), the spread in the coefficients seems tolerable, given the crude nature of the measure being employed. Among the two programs, the Child Care Block Grant produces the narrowest range of coefficients and may, for that reason, be considered the most equitable. Thus, for the purposes of this preliminary analysis, the programs can be rank ordered by equity value, from most to least equitable, as follows:

1. Child Care Block Grant
2. Child Welfare B
3. Child Care AFDC

**Equity Curves**

These findings are portrayed graphically in Figures 1 and 2 with the rough equivalent of a Lorenz curve for each program's distribution. The Lorenz curve illustrates the U.S. income distribution by graphing the fraction of total income received by each percentile of the population (Young 1994, 100–101). The modification employed here graphs the fraction of total federal spending for program X that is received by each percentile of the population of poor children. The curves assign to all poor children in state X the fraction of funds received by that state. The diagonal line in each Figure (labeled "Baseline") illustrates the ideal distribution of strict equality in which, for every state, \( E_x = 1 \). The Figures show that the two most equitable programs come very close to the baseline distribution. The bottom two, least equitable programs produce curves that more closely resemble the income distribution itself, with most of the resources concentrated at the top. These curves confirm the decision to rank the Child Care AFDC program above the Child Welfare E program on the basis of equity.
For those who take egalitarianism seriously in the arena of federal aid to states, these results are disappointing. Although this study did not include an aggregate measure of need-based equity across the four programs, one can make an educated guess about what such a calculation would show. The two most egalitarian programs distributed a total of $950 million in 1991; the two least egalitarian programs distributed more than twice that, or $2 billion. Thus, even if the more equitable programs offset the effect of the less equitable programs, the net effect clearly must be in the direction of greater inequity. Future analyses of other federal programs may find countervailing forces. But for the four programs studied here, the aggregate finding of a residency penalty is clear.

And the Winner Is . . .

Which distributive principle prevails? The analysis suggests that the relationship between the principles of need and contribution operates differently under different funding conditions.

As expected, the one program whose formula was based strictly on need, with no contribution requirement, proved most equitable.
Equity in Child Care
Federal Dollars to States

% Federal Dollars

% Poor Children

when measured against a standard of beneficiary need (Child Care Block Grant). The surprising results come when we compare across the three remaining programs, all of which used formulas that combined measures of beneficiary need, taxpayer need, and contribution. The first of these, Child Welfare B, performed well against the standard of beneficiary need despite the fact that, unlike the other programs, its services are not restricted by law to poor clients. In the last two program distributions, the principle of distribution proportional to beneficiary need appears almost nonexistent. Measured against an equity standard based on child poverty, these programs appear almost arbitrary in their allocations.

These findings may be explained in two ways. First, perhaps per capita income is a total failure as a measure of beneficiary need. The two most inequitable programs examined above relied solely on per capita income as a measure of beneficiary need, whereas the two most equitable programs included other measures as well. Further research is needed to sort out the strengths and weakness of per capita income and poverty as measures of the at-risk population.
Second, the evidence presented here points to a new hypothesis about the relationship among distributive principles: that open-ended funding rules leave the door open for the principle of contribution to swamp the principle of need. The chief difference between the two most equitable programs identified here, and the two least equitable, is a spending cap. The Child Care Block Grant and Child Welfare B are capped while AFDC Child Care and Child Welfare E are not. The spending cap seems to function as a brake on inequity. Only under conditions of open-ended funding (i.e., no cap) does the effect of the contribution principle eclipse the need principle and generate large gaps in state coefficients of equity.

**The Open Door**

Further research also is needed to understand the apparent phenomenon whereby open-ended, entitlement funding leads to glaring inequities in distribution, when measured against standards of beneficiary need. However, the analysis above provides some clues to the mechanisms at work. In particular, the evidence presented here suggests the following rule: Under open-ended entitlement funding, states that spend more, get more, almost regardless of need. The open-ended reimbursement programs with state matching requirements, Child Welfare E and Child Care AFDC, strongly favor states that can pay more up front to serve more federally-eligible clients. Although these states' federal reimbursement rates may be lower than those of poorer states, their total share of federal spending can be disproportionately high if they find ways of including large portions of their client populations in the open-ended federal programs. As they do so, of course, total federal expenditures for these programs will rise steeply. Federal spending for Child Welfare E, for example, grew almost 400 percent from FY 1982–1991, from $374 million to $1.8 billion. The number of children in foster care nationally grew only 47 percent during the same period.

The shifting of program funds toward states that adopt the spend-more-get-more strategy may be an undesirable and unintended consequence from the perspective of the policymakers who designed the open-ended programs. However, the phenomenon should not be dismissed as a breakdown in implementation. Rather, it should be recognized as a basic, structural feature of open-ended social-services entitlements to states. The ways in
which open-ended funding for state social services inevitably leads to uncontrollable federal spending have been carefully documented. As explained below, the same forces that fuel runaway federal spending also lead to inequities in distribution among states.

**How To Use the Open Door**

States that spend more, get more. But why and how do some states spend more? Three mechanisms seem to be at work: state conditions, state policy, and state administrative practice.

Take the example of the largest outlier from the analysis of coefficients of equity in the Child Welfare E program, New York. As noted above, under the existing allocation rule, New York receives almost four times the amount of federal Child Welfare E dollars that it would under a simple rule based on its proportion of the nation's poor children. Part of this disparity is explained by what a former commissioner of the New York City Human Resources Administration has described as "the devastation children face because of the ... dual ravages of crack/cocaine and AIDS" (Sabol 1994, 125–126). New York City's poor populations have suffered disproportionately from the epidemics of crack/cocaine and AIDS. The tragic result is large numbers of children who are orphaned, abandoned, or in need of protection from abusive and neglectful parents. These horrendous conditions help to explain the fact that New York State places a far higher proportion of its AFDC children in foster care than any other state (5.5 percent compared to the national average of 1.7 percent).

State policy in New York also contributes to its larger claim of Child Welfare E funds. New York's basic monthly maintenance rate for children in foster care is the seventh highest in the country. More important, New York is one of a small number of states that pays the full maintenance rate, rather than a lower rate, to relatives who care for children in state custody. Half of New York City's foster children were living with relatives in 1992.

Finally, the disproportionately large allocations of Child Welfare E funds to states like New York seem to be the result of those states' more sophisticated accounting practices (U.S. House of Representatives 1992, 926). For example, some states have developed more efficient methods of identifying foster children who were eligible for, but not actually receiving, AFDC payments at the time of
placement. These children qualify for federal reimbursement under Child Welfare E. One research and advocacy group estimated that states could increase their federal AFDC foster-care reimbursement by as much as $3,000 per eligible child by changing various administrative practices (Center for the Study of Social Policy 1988, 9, 14). These techniques are entirely legal. However, the dramatic variation in their application by states exacerbates interstate inequity.

Similar mechanisms are probably at work in the AFDC Child Care program. States vary greatly in the extent to which they have implemented programs that provide education and training to welfare recipients and the extent to which recipients are working their way off of welfare and taking advantage of transitional daycare benefits. Further research is needed to explain the disparity in state allocations under this program.

Policymakers faced with a choice of capping a social-services program or leaving it open-ended must balance a complex array of concerns. On the one hand, open-ended funding allows greater flexibility for states to respond to social crises that cannot be easily foreseen. On the other hand, open-ended funding creates perverse incentives for states to provide more of the services that are funded in this manner and fewer of the services that are capped (Besharov 1991). It encourages creative accounting practices, which, though entirely legal, do not necessarily lead to better outcomes for children and families. It leads to explosive growth in federal spending for some programs, leaving fewer resources available for other worthy programs. Finally, as this article has shown, open-ended funding generates glaring inequities in the distribution of aid to states.

**CONCLUSION**

Policy analysts generally are not in the habit of evaluating the fairness of the distributions that would result from various policy options. This could change if we start small with programs that distribute federal funds to states. The conditions for analysis are ideal in the case of federal grants-in-aid: the good to be distributed (money) is easily divisible. The claimants to the good (states) are limited in number and fixed in identity. The general principle of distribution, egalitarianism based mostly on need, is widely accepted. Thus, models of analysis should not be hard to develop.
This article has described two promising tools for analyzing needs-based equity in federal grants to states: the coefficient of equity and the variation on the Lorenz curve. These tools should be used to analyze a larger number and a wider array of federal social-services programs, including both existing programs and proposed programs. They should also be used to measure federal funding formulae against other standards of distribution, including taxpayer equity and equity based on contribution. Of particular relevance to policy makers now, these tools can be helpful in measuring the equity effects of block grants as compared with categorical grants and other federal funding vehicles.

In addition, further research is needed to refine and test the methods used here and to develop additional methods for assessing equity in federal distributions to states. Such research might include both model-building and case study research. In particular, case studies of those states, like New York, that represent the outliers in the coefficient of equity tables might shed valuable light on the factors that lead to highly skewed distributions.

When it comes to distributing scarce goods like tax dollars, equity is legitimacy. Thus, issues of equity should not be left as afterthoughts in policy analysis. Policy analysts need to develop new ways of exploring the fairness of distributions.

Notes
1 For more information on all of these programs, see any recent edition of the annual Green Book, published by the U.S. House of Representatives, Committee on Ways and Means.
2 All figures in this article include appropriations to states only, not territories.
3 For a good discussion of the concepts of beneficiary and taxpayer equity, see U.S. GAO (1992).
4 Under Child Welfare E (AFDC Foster Care Program), the costs of maintaining foster-care placements are reimbursed at the Medicaid rate. State administrative costs are reimbursed at a flat rate of 50 percent and state training costs at a flat rate of 75 percent.
5 On the subject of evaluating distributive systems based on the types of information considered relevant to the concept of justice, see Sen (1992).
6 Receipt of federal funds often depends on state compliance with various federal requirements. It would take a leap of faith in the "technology" of federal requirements, however, to equate regulatory compliance with effectiveness in promoting child and family welfare.
For years Senator Daniel Patrick Moynihan has promoted the idea of developing new sets of national social indicators, like the inflation and unemployment rates, that would give us regular, reliable information about the condition of American families. For an excellent summary of existing tools for assessing child and family well-being and their shortcomings, see Zill (1992). The U.S. GAO study (1992) used this disaggregated approach to look at beneficiary and taxpayer equity, but not equity based on contribution. Note that 70–90 percent of all child-welfare cases nationwide are also AFDC cases (Berlin 1992, 113).

Other possible denominators to test include: the number of children who fall below each state's defined need standard, a measure may capture differences in states' cost of living but will also reflect state policy differences and thus complicate interpretation; the number of children below 150 percent of poverty; the number of children eligible for AFDC (also a policy-dependent measure); the number of children eligible for Medicaid (a larger pool, but also policy-dependent).

Ragan and Slottje (1989) used a similar technique to examine the effect of using different definitions of unemployment in federal funding formulae for employment and training grants to states. This effect could reflect higher per capita costs for smaller states setting up programs.

Martha Derthick (1975) tells the story of how the public welfare amendments of 1962 led to runaway federal spending for social services in the late 60s and early 70s (which eventually led to establishment of the capped Social Services Block Grant). Douglas J. Besharov (1991) describes how the same process has caused the title IV-E AFDC Foster Care Program budget to explode in recent years.

References


Governments in many predominantly Muslim states — including Pakistan, Malaysia, and Uzbekistan—have manipulated Islamic symbols and institutions in order to create a national civic identity that would unify their populations' many ethnic and cultural groups. Underlying this objective, however, is the ruling elites' desire to maintain their hold on state power. As the regimes used Islam more and more to further their own interests and those of their primordial or ethnic groups, and as their commitment to Islam became increasingly superficial and regulatory, Islamic opposition movements gained popularity among disenchanted populations, a phenomenon which drove both governments and opposition parties to call for the implementation of Islamic law. Unless the leaders of these states create a cultural or historical value system around which diverse elements of their populations can coalesce, Islam will continue to exacerbate social and ethnic tensions rather than serve as a unifying value for their societies.

*M.A. Candidate, Fletcher School of Law and Diplomacy; B.A., University of Pennsylvania.*
Governments in many Muslim countries have, since they achieved independence, increasingly appealed to Islam to provide the core values, organizing principles, and law of the society. At the same time, many of these countries have witnessed a remarkable growth of civic organizations and opposition political parties that also base their appeal on Islam. What, other than the personal faith of the movement’s leaders and followers, accounts for the growing profile of religion in the politics of the Muslim world?

Certainly many factors are involved, from the failure of Arab nationalism and pan-Islamic ideology, to economic and political crises that afflict many Middle Eastern and Asian regimes. But equally important is the crisis of identity that many Muslim countries faced during and after their struggles for independence. Governments, eager to create loyalty to a newly independent state, often direct the development of a sense of national community. Many ruling groups, such as those of Pakistan and Malaysia, chose to stress Islam as the unifying element of their new nationalism, perhaps demonstrating the veracity of Libyan Colonel Muammar Qaddafi’s assertion that religion and nationalism are the two primary driving forces of human society (Ayoub 1987, 66). In other countries, such as the former Soviet republic of Uzbekistan, the government has emphasized the historic virtues of the ancient Uzbek nation as the core of the new civic identity. Because Islam is still a strong cultural force among the majority of the population, the government has nevertheless found it politically expedient to make concessions to religious interests in order to maintain a veneer of Islamic legitimacy.

Appealing to or avoiding Islam as the foundation for a new national identity is, in these countries, a strategy to help the government elites acquire or maintain control of the state apparatus. “Nationalism is,” as John Breuilly writes, “above and beyond all else, about politics and . . . politics is about power. Power in the modern world is principally about control of the state. The central task is to relate nationalism to the objectives of obtaining and using state power” (Breuilly 1993, 1).

Using Malaysia, Pakistan, and Uzbekistan as case studies, this paper will show that, where governments have used Islam as the core of their “official nationalisms,” they have done so to legitimize the regime in its struggles with competing ethnic or religious power
bases and thereby to maintain their grip on power. This politicization of Islam has exacerbated existing social tensions, as Islamization and the pursuit of Islamic values have become code words for domination by particular ethnic groups—Punjabis in Pakistan and ethnic Malays in Malaysia. Implementing merely symbolic Islamization measures to gain political legitimacy, however, also serves to demonstrate the vacuity of the government’s religious commitment, thereby strengthening the hand of more extreme Islamist opposition forces, who believe that their countries should be governed according to the *shari’ah* (Islamic law).

When the state attempts to create a national identity that downplays Islam for political reasons, as the Uzbek regime is currently doing, it risks basing its legitimacy on an overly intellectualized historical identity that is not shared by the majority of the population. By creating an artificial national identity to serve the political interests of a narrow elite group, the state alienates itself from the population it seeks to influence and thus inspires increased opposition to the regime.

Because the primary objective of Islamic nationalism in all three states has been to keep the regime in power and promote the interests of the ruling elite’s primordial group, official nationalisms based on Islam have both exacerbated existing social tensions and failed to create a sense of civic identity among the population.

**Theoretical Views of Nationalism**

For many individuals in developing countries, even in this age of instant and global communications, the society with which they are concerned extends no further than their village or extended family. As Benedict Anderson explains, this isolation means that “all communities larger than primordial villages of face-to-face contact (and perhaps even these) are imagined . . . because the members of even the smallest nation will never know most of their fellow-members. . . . Yet in the minds of each lives the image of their communion” (Anderson 1983, 6). The social bond between the members of these “imagined communities” is portrayed as a linkage that is strong enough to overcome the myriad other divisions that exist within the society. “Regardless of the actual inequality and exploitation that may prevail,” writes Anderson, “the nation is always conceived as a deep, horizontal comradeship” (Anderson
Islam, which conceives of a worldwide umma, or community of believers, would certainly seem capable of serving as the basis for such global horizontal comradeship.

Islam is at once a religion, a code of conduct for both ordinary individuals and for rulers, and a source of law. It therefore plays a strong role in regulating both social and political structures. Ruling elites can thus emphasize, through state policy, different aspects of Islam as it best furthers their interests and those of the economic or ethnic group that they represent. The leap from culture to politics, writes John Breuilly, "is made by portraying the nation at one moment as a cultural community and at another as a political community" in order to win the greatest degree of legitimacy for the nationalist movement (Breuilly 1993, 69). At different times and under different regimes, Islam can be both (or either) a vaguely defined cultural value or the foundation of the legal system.

Post-colonial ruling groups found that they needed to create a new national civic identity, characterized by loyalty to the central state, that would overcome the wide range of ethnic and regional loyalties that divided their populations (Anderson 1983, 99). They needed to "transform 'primordial sentiments' into 'civil sentiments'" (Jahan 1972, 2).

In many newly independent countries, Islam was the most compelling identifying feature shared by a large segment of the population and thus became the core of these states' official nationalisms. Yet these countries' political leaders refer only to Islamic symbols, since restructuring the entire polity to make it conform to Islamic law and tradition would undermine their personal political interests and those of their primordial group. Instead, they legitimate themselves with Islamic rhetoric that is too vaguely defined to render them accountable when specific Islamic provisions remain unachieved. This practice is consistent with John Breuilly's observation that "ideology ... can have a political impact only if presented in simplified forms and embodied in symbols and ceremonials" (Breuilly 1993, 54). Leaders of all three countries thus make symbolic concessions to Islam, creating academic institutes to study the religion or instituting rarely enforced but high-profile hudud punishment laws. They use Islam, in short, to provide themselves with the legitimacy they need to maintain their grip on power.
Pakistan: Islam in a Multiethnic but Almost Entirely Islamic Society

Islam seems to be a logical common bond to stress in Pakistan, given that Pakistan is one of only two countries to have been created as a homeland for a specific religious group (Burki 1991, 1). The constitutions of 1956, 1962, and 1973 all recognize the special status of Islam, acknowledging that it lies at the foundation of Pakistani society (Zia al-Haq, 1982, 273). The way in which successive regimes called for Islam to enter politics, however, proved to be more divisive than unifying. The forces that drive Pakistani society are based more on ethnicity, family, and region than on religion. Thus, when the ruling Punjabi elite imposed its version of Islam on the entire population in order to reinforce its control over the state and its resources, most people went about their lives—according to their Baluchi, Sindhi, or other ethnically based traditions—as usual.

During the struggle for Indian independence from Britain, Muhammad Ali Jinnah and the Muslim League were primarily concerned with the preservation of Muslim rights within a Hindu majority state. When they concluded that Muslims, who had been treated as second-class citizens under the British Raj, would fare no better under a Hindu-dominated India, they harnessed religious symbols to gain support for an independent Muslim homeland. Religion thus served as "the identifying and differentiating factor" to justify the creation of Pakistan. As a result, the new Pakistani government had little choice but to stress Islam as the core value of their new society. As Edward Mortimer explains:

Muslims were not grouped together geographically, nor did they all speak the same language. If they were not part of an Indian nation, there was no obvious reason why they should consider themselves a single nation at all, unless Islam were to be reasserted as a worldwide nation (umma). (Mortimer 1982, 190)

Although the Muslim League used religious symbols to mobilize Muslims during the independence movement, they did not specify how Islam would manifest itself in the new state's political structures and policies. They had no clear vision of how Pakistan's "Islamic character [would] be reflected in the ideology and institutions of the state" (Esposito 1980, 142). Jinnah and Pakistan's founding fathers did not envision a state where shari'a was the law
of the land. They hoped instead to create a Muslim homeland “in order to enable the Muslims to lead their lives in accordance with Islam. . . . To enable, but not to compel” (Mortimer 1982, 201). They viewed Pakistan as a modern secular state, which would be driven by and based on Islamic values and ideals.

General Ayub Khan, who came to power in a 1958 coup, used religious institutions to consolidate his power. He removed the ulama (clergy) from positions of political influence and appointed lay people to lead the Commission on Marriage and Family Laws and other councils, institutes, and organizations that were charged with defining the nation’s Islamic character (Esposito 1980, 147). Ayub sought legitimacy by creating and expanding Islamic institutions, but simultaneously sought to undermine them to prevent religious forces from becoming independent power bases that might challenge his control. Government concessions to Islam remained equally calculated early in the tenure of Zulfikar Ali Bhutto, who was elected prime minister in December 1970. He and his Pakistan People’s Party (PPP) promised a combination of Islamic values and socioeconomic equality that he called Islamic Socialism, using Islamic slogans and symbols to appeal to rural elites and the urban working class (Weiss 1986, 8). While campaigning, however, he announced no concrete plan for using Islam to restructure Pakistani society. Bhutto intended Islamic Socialism “to be so ambiguous that he could not be criticized by either section of his constituency” (Weiss 1986, 9). His appeals to Islam were sufficient to help him win the election, despite the fact that 113 of the country’s ulama, believing socialism in any form to be foreign and un-Islamic, issued a fatwa (religious decree) stating that socialism is repugnant to Islam (Mortimer 1982, 213–14).

Despite Bhutto’s victory, the 1970 elections should have been a sign to Pakistani political leaders that using Islam as a political instrument was not endearing the government to its population. The three main religious parties received only 13.1 percent of West Pakistani parliamentary seats, as compared to 58.7 percent for Bhutto’s PPP. Such a resounding defeat for the Islamists demonstrated that the electorate rejected a system of government based on Islamic legalism.

The voter remained singularly unmoved by the promise of Islamizing the polity. It seems that he preferred to be left free to practice Islam according to his own inclination and custom. . . . Bhutto’s victory in
the 1970 election must be viewed ... as a popular rejection of the ulama's call for an Islamic state. (Syed 1982, 123)

Ironically, "Bhutto's own exploitation of Islam for political propaganda may also have helped ... to strengthen the credibility of Islamic opposition groups, since it made him more vulnerable to their charges of hypocrisy" (Mortimer 1982, 219). He was particularly vulnerable to accusations that his neglect of Islam led to the secession of East Pakistan in 1971 (Faruki 1987, 57). By 1977, nine religious parties formed a coalition, the Pakistan National Alliance (PNA), to challenge him in that year's elections, forcing Bhutto to address its criticism that the regime had done little to further the cause of Islam. Both Bhutto and the PNA felt the need to be seen as more committed to an Islamic government than the other. Bhutto again made a number of symbolic concessions to legalistic Islam as the election approached, outlawing alcohol and gambling, and announcing in April 1977 that the shari'a would be implemented within six months (Esposito 1980, 151-52). Although a modernizer, Bhutto manipulated Islam's political applications in order to gain legitimacy. "Bhutto could be a Pakistani nationalist as well as a Muslim nationalist, depending on the situation" (Esposito 1980, 132).

Bhutto was not enough of a Muslim nationalist for General Zia al-Haq, who overthrew him on July 5, 1977, and later had him hanged. In his first speech after taking power, Zia stated his commitment to Islam:

Pakistan, which was created in the name of Islam, will survive only if it sticks to Islam. That is why I consider the introduction of an Islamic system as an essential prerequisite for the country. . . . Pakistan is . . . an ideological state. . . . Take Islam out of Pakistan and make it a secular state; it would collapse. (Burki 1991, 69-70)

For obvious symbolic reasons, Zia officially announced the adoption of shari'a as the state law on the Prophet's birthday, February 10, 1979. The legalistic Islamization measures which Zia implemented, however, raised more contentious issues than it solved. The most high-profile laws put in place were the strict budud punishments for sins such as consumption of alcohol, theft, adultery, and other crimes specified in the Koran. In practice, these punishments were rarely applied. Villagers avoided budud punishments by avoiding the court system entirely and using clan leaders to settle disputes. When budud measures were enforced, they were
more likely to be applied to the poor than to the upper and middle classes (Faruki 1987, 63). Zia never adapted Islamic laws to life in modern Pakistan; punishments for robbery were never extended to official corruption, and taxes on camels were never applied to cars and trucks (Faruki 1987, 62). As a result, Islamic punishments failed to either regulate the population’s behaviors or transfer their loyalties to the central state.

Zia tried to extend the state’s reach to even the most spiritual and personal elements of Islam. In August 1984, Zia announced that every village was to designate a person to lead prayers each day, stating that “we are determined to succeed in establishing the system of prayer at all cost” (Faruki 1987, 61). Not surprisingly, such legislation had virtually no impact on the majority of Pakistanis living far from centers of government control. For villagers, Islam is merely

an expression of identity, an idiom of and for morality, and a source of legitimacy. It is viewed as an object, a core symbol, and not as a specific theology, philosophy, body of practice, set of moral injunctions or legal doctrines, and certainly not as a social order promulgated by the current government. (Kurin 1986, 120–21, 127)

The Zia regime continued and expanded its predecessors’ practice of imposing its own definition of Pakistani identity on a society that it neither represented nor understood. Its own conception of Islam demonstrated, according to Pakistani Supreme Court Advocate Kemal Faruki, that the regime confused “Islamization with indigenization” (Faruki 1987, 74). In 1981, Zia’s Punjabi-dominated government made the Punjabi traditional outfit of a shalwar or churidar worn with kamiz or kurta shirts obligatory for all government employees. He later made Urdu, which is spoken by only seven percent of Pakistanis as their primary language (Government of Pakistan 1985, 52), the official language of instruction in Pakistani schools. These measures served as a “highly visible way for the government to combine Islam and nationalism” (Weiss 1986, 12), even though neither the Urdu language nor the Punjabi traditional dress have anything to do with Islam. The state’s efforts to equate Punjabi culture with Islam and portray Punjabi culture as “more Islamic” than other ethnic-groups’ traditions only served to alienate non-Punjabis and non-Urdu speakers and further divide Pakistan along ethnic and regional lines; these policies thus undermined the goals of creating a national civic identity and of emphasizing Muslim unity (Faruki 1987, 74).
Even in a state that was founded as a Muslim homeland, Islam has proven insufficient to provide a national, Pakistani identity that takes precedence over localized primordial ties. This is not entirely surprising given the importance of ethnicity and regionalism in Pakistani society, which is a diverse mix of Islam, saint worship, family relations, and other traditions, all modified by centuries of practice that differ from region to region. As a result, Syed rejects the notion that Pakistani culture is even Islamic, stating that it only has "an Islamic crust" (Syed 1982, 102). It is unlikely that a Punjabi merchant and a Sindhi peasant "see their performance of Islamic duties as having anything to do with issues of national solidarity" (Syed 1982, 102).

Even the regime's symbolic Islamic gestures were rejected by the people. Islamization measures either failed to affect much of the population or, where they did, the population saw through them as hollow promises and as elements in the political struggles between the ruling elites and their opponents. The issue between Muslim nationhood and Pakistani nationhood is not a contest between the Islamic and anti-Islamic sides. It is an issue between the religious establishment and lay politicians; within the second group, it is an issue between those who use the name of Islam to keep their privilege and their challengers. (Syed 1982, 196)

Zia and his predecessors failed to convince the populace that their Islamization measures amounted to more than asserting Punjabi dominance over the rest of the country—a notion highlighted by the secession of ethnically distinct fellow Muslims in East Pakistan. Had the regime emphasized general Islamic values rather than legalistic regulations, Islamization measures might have been seen as more credible, thereby enabling the state to use Islam to transcend ethnic and regional divisions and create a Pakistani civic identity.

**Malaysia: Islam in a Multiethnic, Religiously Divided Society**

In Malaysia, ethnic and religious distinctions run so deep that it would seem ludicrous for the government to place Islam at the core of a new national civic identity, as was attempted in religiously homogenous Pakistan. The newly independent Malaysian government nevertheless attempted to build a national identity based primarily on Malay ethnic identity, of which Islam was considered an integral part.
Due to the growing popularity of Islamist opposition parties, the ruling party ultimately embarked on a process of Islamization out of political necessity. As in Pakistan, the ruling elite of Malaysia turned to Islam to win popular support from its political opponents, using Islam as a legitimating tool to retain power. In Malaysia, however, the competition to appear more committed to Islam than one's opponent has continued to the point that Islamization now threatens the country's fragile ethnic balance.

In 1880, Malaysia was extremely homogeneous. Ninety percent of the population was Muslim and ethnically Malay. The economic needs of the British Empire, however, led to an enormous transformation in Malaysia's demographics. As the British brought Indians and Chinese to work in mines and on railroads, the population swiftly diversified. Malays comprised only two-thirds of the population by 1890 and 49.8 percent at independence in 1963 (Islam 1988, 66–67). Each ethnic group spoke a different language, practiced a different religion, settled in different regions and tended to enter professions and economic strata in which it already predominated.

The primordial divisions that pervade Malaysia's population, particularly religion and ethnicity, reinforce each other—98.7 percent of ethnic Malays are Muslim, 100 percent of Chinese adhere to Chinese traditional religions (primarily Confucianism and Buddhism), and 99 percent of Indians are Hindus (Mutalib 1993, 107). This association affects virtually every element of Malaysian politics. "The very pluralism of the system," explains political scientist Fred von der Mehden, "has led to a society in which ethnicity and religion have become intimately intertwined and in which social policies, politics and economics are all heavily influenced by communal considerations" (von der Mehden 1987, 177).

During the colonial period, anticolonial nationalists represented primarily the poorer ethnic Malays. As a result, they championed Islam to build support for their struggle. Islam and Malay nationalism were, during this period and in the anticolonial context, synonymous and non-conflictual (Abu Bakar 1987, 157). After independence, however, increasing modernization sent many ethnic Malays to cities as poor migrants; in the face of such social dislocation, many Malays developed a greater consciousness of their religious and ethnic distinctiveness (von der Mehden 1987, 193, 197). To alleviate the effects of this demographic transforma-
tion, the Malay-dominated UMNO (Malay Muslim party) government pursued policies that would raise the standard of living of ethnic Malays, which in turn resulted in a still "greater affirmation of Malay cultural identity and in a greater emphasis on religion and ethnicity in Malay politics" (Esposito 1987, 23).

Although these "affirmative action" policies for ethnic Malays were adopted to achieve specific political and socioeconomic goals, the state’s promotion of Islam was inevitably viewed as a threat by other ethnic and religious groups. Nevertheless, the ruling UMNO has tried to improve the position of Muslim Malays without threatening the non-Muslim population’s traditions and rights (von der Mehden 1987, 183). It attempted to mold, in other words, a nationalistic ideology, based primarily on ethnicity, that espouses a Malaysia that is explicitly for ethnic Malays but not necessarily against anyone else.

Both Islam and Malay ethnicity, however, occupy a privileged position in Malaysian politics. Article 160(2) of the constitution requires an individual to be a Muslim (as well as to speak Malay and observe Malay traditions) to be considered a “Malay” and be eligible for the economic and political privileges that go along with such an identity (Esposito 1987, 68). The constitution also requires the head of state to be a Muslim, thus virtually prohibiting any non-Malay from achieving this office (Esposito 1987, 17).

The government was forced to place a yet greater emphasis on Islamic doctrine as several Islamist opposition parties—the Parti Islam Se-Malaysia (PAS) and the Muslim Youth Movement of Malaysia (ABIM)—gained popular support (Anwar 1987, 2). Since the Malay ethnic nationalism of the UMNO calls for allegiance to the Malay nation and not to Islam, the Islamists see the UMNO’s approach to nation-building as un-Islamic and thus illegitimate (Abu Bakar 1987, 155, 163). As the Islamists became more popular, the UMNO “had no option but to try and [sic] co-opt its critics and opponents by appealing to Islam for legitimacy” (Anwar 1987, 91). Put another way, the UMNO “could not afford to leave Islam to the opposition” (Mauzy and Milne 1986, 90).

The UMNO thus embarked on a campaign of Islamization that was similar in content to Pakistan’s. Prime Minister Datuk Seri Mohammad Mahathir announced a range of programs designed to expand the reach of Islam to the population: establishing an Islamic bank, an Islamic university, an Islamic insurance company, and a
Dakwah Foundation to coordinate Islamic education and outreach activities; banning non-halal (ritually slaughtered) beef; making education in Islamic civilization mandatory (later reversed); and considering adopting Islamic morality laws that would apply to Muslims and non-Muslims alike.

Mahathir's appeals to Islam have become so extensive and far-reaching that non-Muslims "no longer viewed [such measures] as concessions [to the Islamists], but rather as part of a government-sponsored Islamization process" (Mauzy and Milne 1986, 92–95; Mutalib 1993, 30–31). Two former prime ministers called for an end to the process in 1983, claiming that Islamization violated understandings with the Chinese and Indian communities that Malaysia would never become a state governed by Islamic law (Mauzy and Milne 1986, 87; Mutalib 1993, 79).

The Islamists, in contrast, were strengthened by government appeals to Islam, which only "provided more ammunition for PAS and others in the opposition to attack the government. PAS views [the state’s actions] as ‘cosmetic Islamization,’ a slow process aimed at momentarily pacifying and containing Islamic revivalism and ensuring that Malaysia never becomes an Islamic state" (Anwar 1987, 92). Even though appeals to Islam threaten the non-Muslim segments of the population, "the Islamization process will not satisfy the fundamentalist counter-elites until it has progressed so far that only they, and not the secular elite, are eligible and qualified to rule" (Mauzy and Milne 1986, 98).

Although Islam served as a unifying political force for the ethnic Malays who benefited from Muslim "affirmative action" programs, the subsequent manipulation of Islam for political purposes has hindered the development of a national civic identity. By creating a rift between Muslims and non-Muslims, exploitation of Islamic ideals endangers the government’s efforts to create a national civic identity based on other grounds. In Malaysia’s multiethnic society, politicized Islam has proven to be more of a divisive factor than a unifying one.

Uzbekistan: Islam in a Society Where the Government Fears It

The emirate of Bukhara and the khanates of Khiva and Kokand, all centered in modern-day Uzbekistan, were long ago among the most important centers of Islamic civilization, acting as centers of trade, science and literature. Although eighty years of Soviet official
atheism virtually eliminated Islam from public view, Islamic rituals and social institutions continued to structure the average person's life. It seems logical, therefore, that in a newly-independent Uzbek Republic, the state would attempt to base its legitimacy on a revival of Islam.

The Uzbek government, led by the same official who headed the republic under the Soviet Union, has indeed used Islamic symbols to boost the legitimacy of the regime. However, because Islamic radicalism has contributed to the civil wars in neighboring Tajikistan and Afghanistan, the government of President Islam Karimov has deliberately downplayed Islam, viewing it as a potential threat to his regime. He has instead chosen to try to unite a diverse mosaic of ethnic groups by appealing to the glorious history of the Uzbek empires. It seems likely, however, that Uzbek society, like that of Malaysia and Pakistan, will end up defining itself from the bottom-up, rejecting the conception of their identities imposed upon them by the government.

Central Asian Islam is a far cry from what is practiced in the rest of the Muslim world. It is infused with saint worship and other pre-Islamic steppe traditions, and peasants say prayers over meals that include vodka. Uzbeks still observe much of Islamic ritual, but they have lost the understanding of the religious doctrine that lies at the root of their practices. Such a heterodox Islam, however, continues in the Central Asian tradition. Even in the fifteenth century Timurid empire, the *Yasa* (traditional steppe law) combined with Islamic practices to structure the everyday lives of the people. Yet even then—as today—the ruler invoked the *shari'a* to provide himself with the legitimacy of Islam (Gross 1992, 16).

Although the Soviet Union outlawed religion, the traditions of everyday popular Islam persevered. Uzbeks continued to hold religious ceremonies for circumcisions, weddings and burials; sanction second wives if a first wife turned out to be infertile; and teach religion to their children (Poliakov 1992, 17–19). The Soviets even legitimized the traditional structures that kept Islam alive by using neighborhood *mahalla* committees to extend their control to the community level. The Soviet empire thus enabled the communities to retain the organizational structure that maintained mosques, taught religion and provided for the collection and distribution of *zakat*. Despite the outlawing of religion (with the exception of a few tightly controlled "official" mosques and religious academies), popular Islam thus had a vast store of resources to keep its traditions
alive. "The mosque (whether official or not) is primarily a social institution that influences all aspects of life and in many ways regulates all of daily life in the mahalla and the family" (Poliakov 1992, 99, 107).

Although Islam was one of the primary defining characteristics of Uzbek national identity under the Soviet Union, Uzbekistan and the other republics surged to independence on a wave of ethnically-based anti-imperialism, reacting to nearly a century of glass ceilings and limited economic opportunities. Yet Islam was an integral part of this newly mobilized Uzbek ethnicity; a December 1988 event to honor the Uzbek language developed into a mass demonstration featuring green flags, signs with Koranic verses, and other Islamic symbols. Similarly, riots in the Fergana Valley in June 1989 were the result of social and economic frustration fueled by a cultural nationalism that synthesized both Uzbek and Muslim identities. The Fergana youths "evinced special enthusiasm not so much for official Islam," writes Central Asian scholar James Critchlow, "but for Timur and the Kurbashis of the 1920s . . . who fought under the banner of Islam" (Critchlow 1991, 176, 180).

As in Malaysia, the mobilization of Islam as an element of a national identity does not concern either non-Muslims or the regime. What scares them, however, is the notion that Islamists will use religion to impose Islamic laws on the entire population. Although many ethnic Russians are emigrating because the resurgence of Uzbek ethnic nationalism has limited their opportunities, "the threat of Islamization is one of the main reasons for the exodus of the Russian population" ("Uzbekistan: Tamerlane v. Marx" 1994, 48).

The government, sensing that Islamic ideology could indeed destabilize the government and divide the country along religious lines (and even along fundamentalist/traditionalist lines), has embarked on a campaign to promote Uzbeks’ historic heritage as the core element of the new civic identity. A monument to Karl Marx in one of Tashkent’s main squares was replaced with one of Emir Timur (Tamerlane), and the observatory of Timur’s grandson, the astrologer-emir Ulugbek, has also been restored in an effort to revive the image of Uzbekistan’s historic and scientific greatness.

Although many Uzbeks speak Russian, President Karimov has manipulated language policy to serve his goal of social and political stability. His regime made Uzbek the official language, despite the fact that President Karimov himself speaks it with middling profi-
ciency. Furthermore, when the government announced that it would replace the Cyrillic alphabet by the year 2000, it decided not to switch to Arabic lettering—which was used before Russian domination and would thus give Uzbeks greater access to their pre-Soviet history—because such a move would enable a freer flow of information from the Arab and Muslim worlds. It instead decided to switch to the Latin alphabet, which will orient Uzbekistan toward the technological secularism of Turkey and the West.

In 1993, Karimov announced that Uzbekistan needed a new ideology to replace Marxism-Leninism—despite the fact that, in order to keep Islam from penetrating the state too deeply, Article 12 of the constitution declares that “no ideology can be considered as the state ideology” (Munavvarov 1994, 144). The objective of the new ideology is “to unite the people around the leader” (Berezovsky 1993, 63), and its primary components are “patriotism and political harmony” (Berezovsky 1994, 83). Maintaining the status quo has, it seems, become the core of Karimov’s social and economic policies, as well as of his attempts to develop a civic identity. “All that we are doing in the sphere of the economy, policy, the social sphere, and the transition to the market,” Karimov told two Moscow reporters, “is connected in principle with the assurance of inter-ethnic social harmony. I am prepared to pay any price for this harmony, for social stability” (Potapov and Biryukov 1993, 79).

Despite his view that an Islamic resurgence would destabilize the country, Karimov still needs to make symbolic gestures that demonstrate his commitment to the faith observed by 89 percent of his subjects (Batalden and Batalden 1993, 167). Because Central Asia has been isolated from the Islamic world, however, Uzbeks are largely ignorant of orthodox Islam, allowing it to be “easily invoked and interpreted in a way that will suit the needs of those in control of the society” (Olcott 1992, xvi). Karimov has therefore treated Islam as more an element of Uzbek culture than as a spiritual commitment; he was thus able to state in the same breath both that the first translation of the Koran into Uzbek “is a great event in the history of our culture” and that fundamentalists are prepared to “use the noble ideas of Islam as a cover to come to power” (“Karimov Fields Questions” 1993, 69).

Karimov has allowed and promoted the construction of mosques and medressehs (religious seminaries). Whereas only 180 official mosques and two medressehs existed in the entire Soviet Union at independence, there are now over 15,000 mosques and 25
medressehs in Uzbekistan alone. He has made Muslim fast days official state holidays, and 3,000 Muslims now make the hajj, the pilgrimage to Mecca, each year. Karimov's personal commitment to Islam, however, remains superficial. An aide who accompanied him on his own hajj admitted that "the greatest difficulty...was abstaining from drinking alcohol and smoking" (Andreyev 1992, 19).

Although he has allowed Islamic institutions to flourish, Karimov carefully controls the ways in which Islam expands. Many mosques have been constructed with the assistance of foreign assistance from Muslim countries, but the regime refuses to accept money for such purposes from Iran, Libya, Saudi Arabia or other "radical" Islamic states. The government allows students to study Islam at such prestigious institutions as Egypt's al-Azhar, but only after it approves of each student. In a blatantly political strategy, Karimov appoints the official religious leadership and then uses it to support his regime; soon after independence, the government-appointed mufti endorsed his patron for the presidency, stating that Karimov is "the most deserving for the now very onerous presidential burden" (Demin 1992, 114).

Yet the concessions that Karimov makes to Islam to boost his legitimacy only provide an environment in which Islam can expand. As happened in Pakistan and Malaysia, Islamic activists will most likely step up their demands on the regime and become increasingly hostile toward a government that attempts to restrict their activities and undermine their growth. Karimov will likely discover, as have the leaders in the two other states examined in this study, that Islam is too deeply embedded in Uzbek society to be surgically removed and replaced with an Uzbek ethnic/civic nationalism that has never before existed; the glorious days of Emir Timur and Ulugbek may inspire intellectuals, but they are far removed from the concerns of Uzbek peasants who grow vegetables in their courtyards to survive. The manipulation of Islam, the core element of most Uzbeks' identities, will very likely backfire on the regime just as such exploitation inspired the radicalization of dakwah groups in Malaysia and exacerbated ethnic tensions in Pakistan.
CONCLUSIONS

The leaders of three states with complex histories, polities and ethnic compositions have tried to change centuries-old traditions of primordial attachments by simply decreeing that they all share a common heritage. Rather than allow the people to make their own conceptions of their identities known to the center, the central governments have informed them of the heritage to which they must feel attached. By using Islam to create a new national identity, the governments of Malaysia and Pakistan have only exacerbated existing social tensions and intensified sources of social and political instability. All three countries have proven the difficulty of replacing cultural norms—especially religiously based traditions and belief systems—with state-imposed values. It is therefore no surprise that the populations of Malaysia, Pakistan and Uzbekistan have either resisted or ignored their states’ efforts, stymieing the governments’ attempts to create a supra-primordial civic identity.

Islam seemed to the ruling elites to be an easy tool to gain legitimacy, manipulate political institutions and retain access to the resources of the state for themselves and for their primordial groups. “Much of the obsession with the shari’a,” it seems, “comes from . . . politicians who depend on these questions of Islamization as a mechanism for seeking access to the state and as a means of reconfiguring alliances when ethnic and regional ties” are in flux as a result of modernization, political change and demographic shifts (Simone 1994, 78). By focusing so heavily on Islamization to legitimize their regimes, however, the governments distracted themselves from the more important tasks of economic development and political institution building, shortcomings that make it even more difficult to maintain stability today.

In using Islam as a tool to gain legitimacy, the rulers of these three countries failed—and even refused—to recognize the complexity of their societies. “Islam is a discursive shortcut,” writes one scholar. “It doesn’t require a great investment of energy or time to serve as a link among people” (Simone 1994, 155). The failure of these three states to successfully build a national civic identity seems to demonstrate that:

it is traditional rural society that will ‘restructure’ the official religious establishment, and not the other way around; nor will the efforts of party officials to derive legitimacy from official religious institutions prove successful in the long term. (Olcott 1992, xxiv)
Politicized Islam drags leaders into a cycle where their commitment to Islam must steadily increase more and more. Eventually, all three states will reach a point where two options exist: (1) they can either implement a wide range of intrusive Islamic measures in order to appease Islamic groups—who will always be able to demonstrate that they are more committed to religion than the incumbents; (2) they can crack down hard on their Islamist opposition, alienating these organizations' supporters. Both options will destabilize the ruling regimes and exacerbate the fragile ethnic and political balances that exist in all three countries unless their governments succeed in finding another system of values and symbols upon which to build a civic identity.

References


CENTRAL BANK REFORM IN LATIN AMERICA: WILL INDEPENDENCE GUARANTEE LOW INFLATION?

John F. Normand*

Since 1989 several Latin American countries have proposed independent central banks as mechanisms for controlling persistently high inflation. This paper examines the questions of political economy surrounding such proposals, with particular emphasis on how four factors—regime type, electoral timing, prevailing economic conditions and the populace's inflation intolerance—contribute to the eventual success or failure of newly autonomous central banks. Case studies on reform efforts in Chile, Argentina, Mexico, and Venezuela suggest that central bank autonomy is more likely to guard price stability when implemented after a period of broader structural reform and inflation stabilization. Attempts to create legally independent central banks in the absence of popular or governmental commitment to price stability may damage institutional credibility in the long run. These conclusions hold potential applicability for other emerging market economies considering similar central bank reform proposals.

A prevailing perspective among policymakers is that low inflation is an essential feature for an efficient market economy and that independent central banks are the best means to achieve this goal.

*M.P.A. Candidate, Woodrow Wilson School of Public & International Affairs, Princeton University; B.A., Georgetown University.
While this belief has guided monetary policy among industrialized countries for decades, the idea of central bank autonomy has only recently gained currency in the developing world. In Latin America the trend toward establishing independent central banks is particularly pronounced, with countries such as Chile, Argentina, and Mexico having recently passed legislation ceding exclusive control over monetary policy to their central banks. Yet despite the momentum toward central bank independence, it remains unclear whether such institutions hold the key to eliminating high inflation in Latin America or other emerging market economies. Most research on the political economy of central bank independence focuses only on industrialized countries and pays little attention to the political context in which central bank reforms are initiated in developing countries.

This paper attempts to illuminate the political dimension by examining several questions of political economy which surround proposals for central bank independence as a means of institutionalizing price stability. Given the high incidence of inflation-prone countries in Latin America, this study focuses on four countries which have recently undertaken central bank reform in the region: Chile, Argentina, Mexico, and Venezuela. Four factors are found to influence the eventual success or failure of newly autonomous central banks. The first factor, “regime type,” refers to the orientation of the government proposing economic reform as either authoritarian, dominant party, or democratic. The second factor, “electoral timing,” tests the hypothesis that central bank reforms are proposed near the end of a chief executive’s term when the ruling party’s loss of power is imminent. “Prevailing economic conditions,” the third factor, refers to the overall macroeconomic climate—especially the inflation rate—at the time of the reform proposal, and tests the claim that central banks are best able to maintain long-term price stability if the government has already lowered inflation to a reasonable level. Finally, “inflationary memory” refers to the collective intolerance of inflation among the populace, which political leaders may harness to generate support for an autonomous central bank.

Three criteria are used for judging a reform attempt as either successful or unsuccessful. First is the willingness of the government to cede monetary authority exclusively to the central bank through legislative reform. Several provisions of reform legislation
are examined, including those concerning authority over monetary policy, term lengths for governors, and conditions for removing governors from office. The second criterion for assessing success is the ability of the new central bank to maintain price stability, as measured by inflation patterns following the promulgation of central bank reform legislation. The third criterion is the management of policy conflicts between the chief executive and the central bank.

The political patterns that emerge from analyzing country experiences in Latin America are somewhat incomplete given the information constraints imposed by the recent timing of reforms. Still, country experiences to date provide evidence to support the following propositions:

1. Legal independence is not an automatic predictor of low inflation in developing countries given the relative youth of central banks in those countries and the lesser attention paid to the rule of law.

2. The current move by high-inflation countries toward autonomous central banks reflects a shift in political culture as embodied in an intolerance of inflation. Legal autonomy is also a stability guarantee to investors and the financial community, which are both politically well-connected interest groups.

3. Successful autonomy plans have been implemented only after price stability has been achieved and are an attempt to limit the policy options of successor governments. Regime type may facilitate reform in some cases, but it is not a necessary condition.

4. Finally, the ultimate success of a central bank in achieving price stability rests on its credibility with consumers, business, labor, and finance. How a government transfers credibility to a newly independent central bank is crucial to overall success. These conclusions, discussed later, hold important policy implications both for countries on the verge of reform consolidation and for those designing a reform program that includes overhauls of the financial sector.

**THE INADEQUACY OF LEGAL INDEPENDENCE ALONE**

Several theories have been offered to explain why independent central banks are more likely to promote price stability than nonindependent ones. The prevailing notion holds that political leaders will use the influence over central banks in order to increase the money supply, stimulate economic growth, and thus win
popular support to guarantee reelection. An independent central bank (i.e., one which cannot be forced to lend directly to the central government or to public sector enterprises) insulates the economy from the political-business cycle because it can resist executive manipulation of monetary policy. The inability of politicians to exercise restraint over inflationary monetary expansion in order to garner popular support has been offered as the most compelling argument for central bank independence (Volcker and Mancera 1991).

Several recent studies support the hypothesis that countries whose central banks exhibit the highest degree of legal independence also demonstrate the lowest rate of inflation (Bade and Parkin 1988; Alesina 1988, 1989; Grilli, Vitorio, and Masciandaro 1989). Cukierman, Webb, and Neyapti (1992) have found that while legal central bank independence is a good predictor of low inflation in industrialized countries, it is not a good predictor in developing countries.

The notion that indices of legal independence hold less predictability for inflation stability in developing countries should come as little surprise given the methodological assumptions inherent in these indices. Most indices are based on written bank charters and specifically exclude considerations of how the law is applied in practice. Given this methodology, these analyses hold less explanatory power in countries exhibiting a high disparity between legal theory and practice—that is, in countries where the rule of law may not be explicitly recognized as a fundamental governing principle. This disparity is particularly pronounced under authoritarian regimes where most political decision making may be concentrated in a small group of political actors or, as likely, in countries undergoing structural economic transformation where power relationships may be in flux. In such cases, ranking the independence of central banks should take into account alternative nonlegal factors such as the presence of a strong central bank president, the competence of the central bank staff, or turnover rate of the central bank governors (Cukierman, Webb, Neyapti 1992, 361).

As Maxfield (1994) notes, rather than rely on the rule of law, central banks in developing countries instead react to the financial needs of government officials and the private sector financial market. The more robust the financial markets, the greater incen-
Tive financiers have to protect their interests by pressuring the central bank and government officials toward price stability. Similarly, the more developed the financial markets, the less dependent the government is on central bank financing and the more likely that it will tolerate an independent central bank.

While Maxfield's hypothesis may explain why countries find themselves more likely to support continued central bank independence and price stability over the short run, her explanation does not clarify the motivations pushing governments to propose alternative central bank arrangements. Should we conclude that only countries with well-developed financial markets would have relatively independent central banks to protect financial interests? Numerous other countries, such as Brazil, Venezuela, and Peru, have active and well-developed financial sectors, a growing share of industrial production in GNP and thus a growing class of industrialists, yet exhibit no more penchant for price stability than Nicaragua or Zambia. Clearly, a more complicated dynamic than financial sector incentives is needed to explain the effectiveness of central banks.

The case studies which follow illustrate some of the additional factors such as regime type, inflationary memory, electoral timing and prevailing macroeconomic conditions that also contribute to the eventual success or failure of a country in establishing an autonomous, inflation-fighting central bank. Table I provides summary statistics on inflation and central bank autonomy indices on Argentina, Chile, Mexico, and Venezuela—the four countries that this paper will use to examine the conditions under which central bank autonomy is most likely to guard price stability.

**Case Studies in Successful Central-Bank Reform: Chile and Argentina**

**Chile**

Chile's central bank as reorganized in 1989 is arguably the most legally independent in the developing world. It is forbidden by constitutional amendment from printing any new money to cover a fiscal shortfall except in wartime. Its five governors are appointed for periods of ten years, with staggered departures every two years, thus giving them the ability to outlast two presidential administra-
Governors cannot be removed for any reason by the chief executive or the Congress (Marshall 1991; Fontaine 1989).

Such an arrangement represented an abrupt about-face for the Banco Central de Chile, which previously could assert little institutional autonomy and acted as a virtual checking account for government deficit financing. Prior to reform, its governing board consisted of members appointed by Chile’s president, the national banks, and private sector representatives from industry and agriculture. Since the 1940s it had been responsible either directly or indirectly for financing the state industrial holding company CORFO and encouraged other commercial banks to do the same. Other organizations involved in agricultural or regional development were similarly authorized to solicit funds from the central bank. This habit continued in subsequent years and peaked when Salvador Allende’s Marxist government came to power in 1970. Consistent with its socialist agenda, the Allende regime authorized an enormous increase in social services and salaries which compounded demands on the central bank for additional government financing. The result was not surprising: inflation shot up from 20 percent at the beginning of Allende’s term in 1970 to over 350 percent at the time of the coup d’etat in 1973 and averaged nearly 175 percent annually for the remainder of the decade (See Table 1).

<table>
<thead>
<tr>
<th></th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>90</th>
<th>91</th>
<th>92</th>
<th>93</th>
<th>Independence Index</th>
<th>Turnover Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>23.5</td>
<td>132.9</td>
<td>565.6</td>
<td>2314</td>
<td>171.7</td>
<td>24.9</td>
<td>10.6</td>
<td>.40</td>
<td>.92</td>
</tr>
<tr>
<td>Chile</td>
<td>30.4</td>
<td>174.5</td>
<td>21.37</td>
<td>26</td>
<td>21.8</td>
<td>15.4</td>
<td>12.7</td>
<td>.46</td>
<td>.46</td>
</tr>
<tr>
<td>Mexico</td>
<td>2.7</td>
<td>14.7</td>
<td>69.1</td>
<td>26.7</td>
<td>22.7</td>
<td>15.5</td>
<td>9.8</td>
<td>.34</td>
<td>.15</td>
</tr>
<tr>
<td>Venezuela</td>
<td>1.7</td>
<td>6.6</td>
<td>23.05</td>
<td>40.8</td>
<td>34.2</td>
<td>31.4</td>
<td>38.1</td>
<td>.43</td>
<td>.30</td>
</tr>
</tbody>
</table>

Upon seizing power in a 1973 coup, General Augusto Pinochet undertook several steps, including central bank reform, to reverse the economy’s inflationary trend. Attempting to reverse what he considered the profligacy and disorder generated by the Allende administration, Pinochet installed a core group of economic advisors trained at the University of Chicago who proceeded to implement what has been called a “radical experiment with neo-conservative economics” (Foxley 1983). The Pinochet government followed several years of inflation stabilization with a 1980 constitutional amendment prohibiting the central bank from printing money to cover fiscal deficits. In 1989 he strengthened these reforms by proposing to make the central bank fully independent of the executive branch and Finance Ministry by lengthening the terms of office of the governors and giving the central bank direct control over monetary and exchange rate policy. The measure was easily implemented through executive decree and became law in 1989.

The success of Chile’s central bank has been evident in recent years. From a strictly legal perspective, the Chilean central bank charter contains the key strengths of the most respected independent central banks in industrialized countries: exclusive control over monetary policy, long tenures for governors exceeding those of the chief executive, and freedom from arbitrary dismissal by the chief executive. In practice, the past five years of autonomy have witnessed few heated contests of will between the chief executive and the central bank. Further, annual inflation has fallen to single digits for the first time in 30 years. The Chilean central bank reforms have even been cited by other governments in the region as models for autonomy.

Explaining both the policy shift in Chile and the success thus far in maintaining price stability requires analysis of the political context of reform. The electoral timing of the policy proposal is the first key element in explaining why Pinochet, who had achieved de facto monetary autonomy through constitutional amendment in 1980, would propose central bank autonomy in 1989. In 1988, after fifteen years as chief executive, Pinochet proposed a plebiscite on his administration under which a “yes” vote would sustain his rule for eight additional years while a “no” vote would automatically trigger a presidential election in 1989 in which opposition candidates would be allowed to run. Pinochet lost this referendum, and
by 1989 it was clear that the Chilean electorate would vote Pinochet out of office. Faced with imminent demise, Pinochet initiated a series of provisions to ensure his stamp on Chilean politics and policy-making would endure through the subsequent regime (Herrara and Graham 1994, 250–53). The Pinochet government’s proposal called not only for a more autonomous central bank, but also gave Pinochet himself the power to appoint all five incoming governors, a power which would allow pinochistas to control monetary policy through the year 2000. Pinochet’s move to grant autonomy thus can be seen as an attempt to institutionalize his own policy preferences for monetary stability well after he or his advisors had left office and to ensure that his successor would not retreat from the program that had brought strong growth to Chile in recent years.

Chile’s authoritarian political structure further facilitated such a bold policy move in the face of opposition. The 1989 proposal by Pinochet was hotly contested among both economists and political leaders. Patricio Aylwin, who defeated Pinochet in the 1989 election, opposed the idea, arguing that it would create a central bank even more powerful than the German Bundesbank or the U.S. Federal Reserve.3 Chilean economists also questioned the soundness of creating a “superbanco.” However, as has been noted with respect to other structural adjustment reforms in Chile, the Pinochet regime had no need to “sell” the reform proposal to the public given its exclusive hold on power (Herrera and Graham 1994). As a dictator writing the terms of his own departure from office, Pinochet had the luxury of writing in provisions which may not have been possible in a more decentralized political context.

This is not to say that Pinochet completely disregarded public opinion. After significant opposition from Aylwin’s newly elected government, Pinochet did concede on his earlier demand to appoint all central bank board members by allowing Aylwin to appoint two members and agreeing on a joint appointment for board president. Further, Pinochet could still garner some measure of public sympathy for his proposal given the country’s inflationary history. As in other countries which had experienced hyperinflation, such as Germany after World War I, Chileans could still recall the hyperinflation of the Allende years. This public intolerance of inflation lent some justification to Pinochet’s policy prescription, even if implemented under a repressive regime. A public consen-
sus on the need for price stability, while not necessary for passing legislation in an authoritarian context, could nonetheless be seen as a way to establish institutional legitimacy for the central bank's role in policy-making during the post-Pinochet years. Therefore, the combination of previous hyperinflation, a strong government committed to anti-inflationary policies, and constitutional prohibitions against deficit financing have all contributed to the record of price stability and institutional credibility that the Chilean central bank enjoys today.

Argentina

In contrast to Chile, which undertook structural adjustment under authoritarian rule, Argentina's reforms progressed under democratic regimes, which have held power since the country's military rulers stepped down in 1983. When Carlos Menem was elected President in 1989, the Argentine economy had reached the apex of its inflationary struggle. Inflation had averaged over 500 percent annually throughout the 1980s and had surpassed 3000 percent by Menem's inauguration. The policies of the Banco de Argentina were a leading source of such price instability. Like other Latin American central banks, the Banco de Argentina subsidized state enterprises and absorbed internal and external public debt throughout the 1980s. By some estimates the central bank had been responsible for nearly $68 billion in losses during the 1980s, an amount equal to Argentina's entire foreign debt or one year of national income (World Bank 1993). The central bank had been forced to lend to other financial institutions with no expectation of repayment and had been pressured to issue treasury bills which it could not cover. Numerous financial demands on the central bank and the lack of alternative financing mechanisms in Argentina left the central bank with a continual need to print money or issue financial liabilities to cover the numerous debts it had assumed. The resulting hyperinflation of the late 1980s was therefore not surprising.

Given the central bank's historically lax supervision of monetary growth, its reform was a fundamental component of any long-term stabilization plan. Upon taking office in 1989, Menem proposed to legally limit the chief executive's influence over monetary policy. The Law of Convertibility was inacted to peg the peso to the U.S. dollar and prohibit the central bank from printing money unless it were backed by foreign-currency reserves, gold, or a limited
quantity of government bonds (World Bank 1986). Two years later, President Menem proposed to render the central bank fully autonomous by giving it exclusive control over the peso, restricting government borrowing from the bank, and removing extraneous responsibilities not directly related to price and exchange rate stability.

Argentina's success with maintaining price stability in recent years has indeed been remarkable. From a high of over 3000 percent in 1989, annual inflation fell to less than eleven percent in 1993 and to single digits in 1994. Further, Argentina's Law of Convertibility has created a monetary regime that extends beyond traditional notions of central bank autonomy and further insulates policy-making from executive pressures. By creating a currency board, the Law of Convertibility restricts the ability of the chief executive to manipulate monetary policy for political gain. At the same time the currency board reduces the discretion of central bankers who might be sympathetic to the chief executive's agenda. Such a system has the additional advantage of creating confidence in the fiscal solvency of the country since reserves must cover the monetary base. This feature can be credited for helping the Argentine peso avoid speculative attacks in the wake of the Mexican peso devaluation of December 1994.

That Argentina has replicated Chile's price stability is particularly interesting given the decidedly different political context under which these two central bank reform efforts were undertaken. Unlike Pinochet who enjoyed dictatorial powers in Chile, Menem faced the challenge of implementing economic reforms within a democratic political structure of more diffuse power relations. Imposing institutional reforms by executive fiat was therefore not a policy option as in Chile. Given that central bank reforms were proposed early in Menem's term, it does not appear that electoral timing played as important a role in influencing the reform effort as it did in Chile.

Lacking the institutional authority or the electoral motivations of Pinochet, Menem thus availed himself of the remaining political resource in Argentina, growing Argentine intolerance for inflation. Menem assumed the presidency in 1989 at a critical juncture when successive stabilization plans had failed and inflation had topped 3000 percent annually. The effects of such a crisis on political culture and interest-group alignments proved profound. Economic crises, such as the one Argentina experienced during the
hyperinflation years of the 1980s, can redirect policymakers out of their traditional modus operandi, realign interest groups that would otherwise oppose drastic policy reform, and create new coalitions to pressure government action (Williamson 1994). Argentina's inflationary crisis of 1989, more acute than Chile's inflationary struggle of the 1970s, required central bank reform measures more drastic and deep-rooted than those implemented in Chile. To break the inflationary expectations of the public, the government had to implement a policy that would prove its resolve to monetary stability and establish the long-term credibility needed to maintain low inflation. Given these exigencies, the Menem administration pushed for a Law of Convertibility, which until that time was the most drastic legal measure by a Latin American government to impose monetary discipline on a central bank. By the time full central bank autonomy was legally established two years later, the Menem administration had not only proven its commitment to price stability, but had imposed a legal safeguard against expansionary monetary policy. Having established its commitment to follow anti-inflationary policies through the convertibility law, Menem's administration could more easily establish a credible central bank capable of sustaining low inflation in subsequent years.

CRISES OF CREDIBILITY: CENTRAL BANK REFORMS IN MEXICO AND VENEZUELA

Mexico and Venezuela serve as useful counterweights to the Argentine and Chilean cases of central bank reform. Both countries have witnessed extreme fluctuations in their economic fortunes over the past decade. Venezuela, praised until the 1980s for its record of high growth with low inflation, was blacklisted in 1994 by international investors due to President Rafael Caldera's decision to re-institute capital and foreign exchange controls, nationalize several banks, and impose price controls on numerous consumer goods. Mexico, considered in the early 1990s to be a model policy reformer among emerging market economies, found itself, in 1995, saddled with a fiscal austerity plan, capital flight, and rising inflation following the December 1994 peso devaluation. The following sections explore how central bank policy has contributed to abrupt increases in inflation in both countries.
By Latin American standards, the Bank of Mexico prior to the 1980s was a relatively autonomous, inflation-fighting central bank. Given the institutional structure of the bank prior to reform in the 1990s, this should indeed seem surprising. Before central bank reform in 1993, the Bank of Mexico set no legal limits on government borrowing from its coffers. Further, the Bank shared decision making on monetary and exchange-rate policy with the Ministry of Finance and the Ministry of Budget and Planning, and it relied heavily on them to resist expansionary pressures from political officials. Despite the potential for abuse of monetary policy under such an arrangement, the Bank in fact presided over a period of relative price stability through much of the 1950s through 1970s (Maxfield 1994).

During the 1980s, however, exogenous economic shocks eroded this commitment to an anti-inflationary monetary policy. Higher international interest rates and lower oil export prices dealt a double blow to the balance of payments and increased the need to seek alternative financing through the central bank. Faced with a growing balance of payments crisis, a lack of financing alternatives, and a host of emerging collapses among commercial banks, then-President José López Portillo nationalized the private banking system, including the Bank of Mexico, which was placed under the Ministry of Finance. With no legal limits to restrain government borrowing from the central bank, the government moved to monetize its deficit. This resulted in inflation that reached as high as 150 percent in 1988 at the election of President Carlos Salinas de Gortari.

Though inflation stabilization was a key feature of Salinas's overall economic-reform agenda, his early initiatives such as the Pact for Economic Solidarity, which bound business, government, and labor to accept price and wage restraints, made no mention of instituting an autonomous central bank. While the issue had been broached by several leading policymakers, including Banco de Mexico Governor Miguel Mancera, a serious proposal to implement such a plan was not floated until 1993. In May 1993 Salinas proposed a constitutional amendment to free the central bank from the Ministry of Finance, under which it had been subordinated since the bank nationalization of 1982. The PRI-controlled Con-
gress approved the amendment in December 1993, and it took effect in April 1994. The amendment established price stability as the main and exclusive priority of the central bank and limited the amount that the Bank of Mexico could lend to the federal government to 1.5 percent of the federal budget. To further insulate governors from executive influence, the amendment set the term of the governor at eight years (exceeding the President's by two years) and staggered terms of deputy governors to avoid presidential board packing.

Until December 1994 it appeared that Mexico's central bank reform was an unqualified success. The liberation of the central bank from the Ministry of Finance and the extension of governors terms to eight years were fundamental reforms, although the staggering of terms still maintained some opportunities for executive influence.\(^5\) In practice the central bank also performed quite well until the December 1994 devaluation. By 1993 inflation had fallen to single digits for the first time in two decades and had remained at this level for almost two years.

That the newly independent central bank had established an initially credible track record was not unexpected given that the reform process in Mexico reflected several aspects of successful reform in Chile and Argentina. Like Pinochet in Chile, Salinas was well-positioned to institutionalize low inflation policies by leveraging his tremendous power as chief executive in a dominant party state. Moreover, Salinas could capitalize on the accumulated intolerance of inflation in Mexico. Through the Pact for Economic Solidarity, Salinas had not only lowered inflation to single digits for the first time in two decades, but also forged societal consensus on the maladies of inflation and the need for restraint on the government's temptation to practice populist policies which the budget could not finance (Kaufman 1994).

Yet despite these factors that argued in favor of a strong inflation-fighting central bank, Mexico found itself on the brink of an inflationary crisis in December 1994 following the peso devaluation. Why did the same factors such as electoral timing or regime type, which favored successful central bank reform in Chile, or the presence of strong anti-inflationary sentiments, which supported drastic institutional reform in Argentina, not also lead to sustainable low inflation in Mexico following the creation of an independent central bank in 1994?
Undoubtedly, some of the inflationary tendencies touched off by the December 1994 devaluation could be attributed to overly pessimistic predictions about the state of the Mexican economy. When investors hold overly pessimistic expectations and fear a devaluation, they will try to convert peso holdings into dollars, which weakens demand for pesos and precipitates a fall in the peso’s value. Thus, fears alone of a devaluation may give rise to the devaluation itself and the concomitant inflationary pressures due to rising import prices.

This overcompensation of the market notwithstanding, Mexico’s monetary crisis was also attributable to the central bank’s monetary policies in the eight months preceding the December 1994 devaluation. Beginning in March 1994, when Mexico entered the thick of presidential campaigning, the central bank increased the money supply by over 20 percent (IMF 1995), presumably with the intention of boosting the PRI’s chances of victory in the upcoming elections. Further, the central bank did not release information on the low level of foreign reserves until after the presidential election, for fear that doing so would force an earlier devaluation and weaken the PRI’s electoral prospects. Such policies suggest that, despite its independent legal status, the Bank of Mexico was not as insulated from executive influence as originally purported. The Bank of Mexico has indeed suffered a severe blow to its institutional credibility. As the case of Venezuela will illustrate next, such a crisis of credibility has long-run consequences that are not easily overcome.

Venezuela

Like many central banks in Latin America, the Banco de Venezuela exhibited few characteristics of legal autonomy. Prior to reform efforts in 1992, it did not exercise exclusive control over formulating monetary policy but instead shared some responsibilities with the government. During conflicts with the executive branch, the central bank did not always have the final word. In one notable incident the government actually sought and obtained a ruling from the Venezuelan Supreme Court ruling to reverse a central bank decision to lift interest rate ceilings. Yet despite these legal deficiencies, Venezuela maintained a remarkably low inflation rate until the 1980s. Hausman (1990) attributes this to the “fiscal deficit aversion” of Venezuelan policymakers whose tendencies to run deficits were
proscribed until the 1980s by legal requirements for a balanced budget and limits on public credit expansion.

This strong anti-inflationary record collapsed in the 1980s when Venezuela confronted the structural adjustment crisis that had plagued other Latin American countries. According to Nairn (1993), the complacency of oil-led growth disguised structural imbalances in the Venezuelan economy which had been developing over years and later precipitated the inflation and fiscal crisis of the 1980s.⁶ Successive stabilization plans had failed, and by the late 1980s the state was facing virtual fiscal insolvency.

In 1989 Venezuelans elected Carlos Andres Perez as President. Although he campaigned on a populist platform, once in office Perez espoused the same deficit reduction and trade liberalization policies of other neoconservative leaders such as President Salinas of Mexico. Perez's reforms were immensely unpopular both among the public (who rioted following a bus-fare increase in 1989, leaving 60 people dead) and the military. Two failed military coups in 1992 and corruption charges were enough to drive Perez from office later that year, but not before the Venezuelan Congress passed legislation giving the central bank control over monetary policy and a floating exchange rate. Perez appointed a new six-person board headed by Ruth de Krivoy, a strong proponent of monetary conservatism, whose five-year term would outlast the chief executive's by one year. However, by stipulating that the president and directors could be removed for a variety of reasons such as "failure to perform the responsibilities of the position," the new law preserved opportunities for governmental interference in monetary policy-making (Financial Times, June 23, 1993).

As restructured in 1992, the Banco de Venezuela could not be considered as successful in maintaining price stability or resisting executive pressures as central banks in Chile or Argentina. In practice, the bank has been unable to preserve low inflation, which rose from 31 percent in 1992 to over 100 percent in 1994. Moreover, the 1992 reform legislation still maintains provisions for unseating governors, a problem that has hindered the bank's ability to resist executive pressures to manipulate monetary policy for political gain. For example, in April 1994 central bank president Krivoy and half the central bank board resigned after newly-elected President Rafael Caldera pressured the bank to lower interest rates. She and her three colleagues have since been replaced, thus giving calderistas a majority on the central bank board.
The failure of Venezuela’s central bank to assert its position as an inflation-fighting force once legally liberated from the executive branch raises several concerns about the essential preconditions for successful central bank reform. Comparisons to Chile, Argentina, and Mexico yield some insights. As in Chile and Mexico, central bank reform in Venezuela was proposed near the end of tenure for a chief executive who had initiated significant policy reform in his country. His efforts can thus reasonably be interpreted as an attempt to institutionalize reforms under his administration and thereby prohibit his successor from retrenching. However, Venezuela’s regime type and prevailing economic conditions further complicated the reform process. Unlike Pinochet or Salinas, who as chief executives in dominant or one-party systems could successfully manage their departures from office, Perez was not empowered to do likewise. The Venezuelan president enjoyed few of the broad institutional powers afforded by the Mexican or Chilean presidencies or the political capital controlled by their respective party apparatus. With less lawmaking discretion than either Salinas or Pinochet, Perez could not push through extensive central bank reform easily. Instead of proposing to further lengthen governors’ terms or to prohibit the chief executive from unseating governors, Perez settled for more modest measures. Electoral position may have provided the opportunity for reform, but regime type did not guarantee the creation of a strong central bank to consolidate reform.

Venezuela also differed from Argentina and Chile in terms of the prevailing economic conditions at the time when central bank reform was proposed. In 1992 when central bank autonomy was first proposed by Perez, inflation still hovered around 30 percent annually. By contrast, Argentina and Chile’s inflation rates were under 20 percent and were declining. Both the Argentine and Chilean governments had not only proven capable of lowering inflation, but had also shown a commitment to policies over a sustained period which kept inflation at low levels. Only after having established this credibility as inflation fighters could the Chilean and Argentine governments then transfer such powers to a central bank which would lock in these reforms. Venezuela’s central bank lacked the benefit of such a “credibility transfer.” Furthermore, President Perez was ousted from office before he could lower inflation significantly and begin the process of consolidation with some measure of credibility.
Venezuela’s experience highlights the notion that legal central bank reform is a necessary but not sufficient condition for institutionalized reform consolidation. Equally important are bank legitimacy and credibility. Legitimacy refers to the notion that policymakers and the public alike acknowledge the importance of price stability and designate the central bank as the most appropriate guarantor of this policy. Credibility flows from legitimacy in that the actions of a central bank will not be taken seriously by the financial community or labor leaders unless those actions are seen as being within its legitimate mandate.

As rational expectations adherents have argued, the credibility of central bank actions are indeed critical for price stability. According to these theories, inflation is partly determined by the expectations of the government and of the public as each tries to forecast the other’s move and act strategically in response. If the public expects the monetary authority to run an inflationary policy, producers will seek higher prices, labor will demand higher wages and investors will ask for higher interest rates. Inflation’s impact is compounded by expectations, which further complicate the stabilization process. Given the importance of expectations, governments try to promote institutional changes such as increasing central bank autonomy that, in addition to exercising tangible control over monetary policy, also alter the psychological elements of inflation by reducing inflationary expectations. According to inertial inflation theorists, a government that seriously espouses low inflation should support central bank independence in an attempt to influence long-term expectations.

Part of the Bank of Venezuela’s weakness is that it lacked sufficient credibility to either influence expectations or resist pressures from the chief executive. The mere promulgation of a new central bank charter was insufficient to influence expectations among consumers or financiers without a commitment from the government to cut the fiscal deficit. Thus inflation persisted despite central bank reforms. Further, the resignation of the central bank governors demonstrated that despite legal reform, the chief executive still influenced monetary policy in Venezuela. Though the resignations of Krivoy and her colleagues were voluntary, a major personnel change in the central bank was rumored from the time Caldera assumed office. Given these limitations, the outcome of Venezuela’s experiment with central bank autonomy are not unexpected.
The experiences of Mexico and Venezuela with inflation stabilization illustrate that policy reform is not always a linear process. Countries that may at one point undertake significant structural overhauls may find their efforts reversed by an administration that does not share their policy preferences or by external developments which jolt a fragile policy infrastructure. Having traversed the political minefield of inflation stabilization, IMF conditionality, and deficit reduction in the 1980s, Latin American countries are finding that an additional obstacle remains, reform consolidation. The notion that policy reforms cannot be self-sustaining without complementary institutional arrangements has proven painfully true in both Venezuela and Mexico.

At the same time, it is clear that institutional reforms that lack political commitment are equally dangerous. Disingenuous attempts at central bank reform by one administration may in fact inhibit the ability of policymakers to establish sustainable reforms in the future by reducing the credibility of bank actions. Such is the challenge faced by the central banks of Venezuela and Mexico.

Given the need for consolidation mechanisms and the potential loss of credibility from half-hearted attempts, policymakers undertaking central bank reforms in Eastern Europe, the former Soviet Union, or elsewhere in Latin America would be well-advised to look toward the experiences of Argentina, Chile, Mexico, and Venezuela. A review of both the theoretical literature and case studies in Latin America yields several conclusions about the relationship between autonomy and price stability and about the political conditions under which central bank autonomy can be an effective means of reform consolidation. First, legal independence is not an automatic predictor of low inflation in developing countries given the relative youth of central banks as institutions in these countries and the lesser weight placed on adherence to the rule of law. Thus, an explanation of central bank behavior must delve more deeply into nonlegal measures of independence and the political context of reform efforts.

Second, even though legal measures of central bank independence are a poor predictor of price stability in developing countries, many governments are nonetheless moving to make their central banks legally independent of the executive branch. This move should not be viewed cynically as a token gesture toward
Central Bank Reform in Latin America

autonomy that will be disregarded when convenient. Rather, the cases of Chile and Argentina illustrate that this policy shift reflects an evolving consensus among policymakers that price stability is a reasonable, if not essential, policy goal. Further, the experiences of the countries studied suggest that reform-minded governments should strengthen institutions to accomplish this end. Intolerance of inflation is becoming an increasingly embedded notion in the political culture of formerly high-inflation countries in Latin America. This changing political culture acts to both encourage policy reform and lend credibility to institutions established to preserve anti-inflationary interests. In addition, public proposals for legal central bank independence also signal investors that the value of their assets will not be eroded by excessively inflationary monetary policies, a critical assurance in an age of increasing capital mobility.

Third, successful reform efforts have been implemented by outgoing regimes which attempted to consolidate broader structural reforms before leaving office. While regime type may have facilitated policy reform in Mexico and Chile, it is not clear that such a structure is a precondition of reform everywhere. In Argentina for example, President Menem successfully leveraged the public's intolerance of inflation to create an independent monetary regime even though he lacked the broad executive powers of Pinochet or Salinas. However, a more in-depth test of the regime type hypothesis will require examining future reform efforts by additional countries with democratic political structures. Lacking absolute presidential powers to push through reform, those leaders will need to use other means of building societal consensus, such as capitalizing on an inflationary crisis or a strong public aversion to inflation. This tactic is particularly important in the emerging democracies of Eastern Europe and the former Soviet Union.

Fourth, successful central bank reform has been a preventive measure to preclude backsliding once inflation has been stabilized rather than a remedial measure to subdue hyperinflation. Legal central bank autonomy is not in itself a sufficient measure for curing inflation. As the case of Venezuela illustrates, mere legal reform without a true government commitment to fiscal stability will prove unsuccessful in the absence of government commitment to deficit reduction. Indeed, it may hinder the credibility of future reform efforts.
Finally, a central bank’s long-term effectiveness in consolidating price stabilization hinges on its credibility. Legal charters granting autonomy are critical in this process, but they are not sufficient, especially in countries where the chief executive is unaccustomed to restrictions on the exercise of power. The best, in fact primordial, condition for bank credibility is inflation stabilization prior to bank independence. Credibility can also be created by a strong leader committed to price stability (Menem or Pinochet), a sustained period of low inflation (Chile), or an appeal to anti-inflationary sentiments among the populace (Chile and Argentina). Without such credibility, attempts to make central banks legally autonomous will prove unsuccessful at securing long-range price stability and reform consolidation.

Notes
1. The potential for conflict of interest is evident from the board’s composition. See Marshall (1989, 41).
2. CORFO (Corporación de Fomento de la Producción), founded in 1939, was partially financed through commercial bank loans made at only a 2 percent annual rate, payable over 50 years. Banks were encouraged to extend such generous terms by the central bank, which counted these loans against the branch banks’ reserve requirements owed to the central bank. The effect was the same as if CORFO had been subsidized directly by the central bank. For a discussion of the relations among CORFO, the central bank, and the industrial sector, see Marshall (1991).
3. Aylwin’s claim was partly motivated by the belief that General Pinochet was planning to hand over too little power to the successor government. See “Chile Plans to Create Strong Central Bank,” New York Times, October 23, 1989, p. D14.
4. For example, the Central Bank routinely covered capital and foreign exchange losses, as well as the outstanding debt from banks liquidated in the financial crisis earlier in the 1980s. See World Bank (1993).
5. The deputy governor’s terms are structured so that by the fourth year of a six-year administration, the President will have appointed three of the five sitting governors, giving him a majority on policy decisions, should he choose to exercise it. In contrast, the U.S. President can name only two of the seven Federal Reserve members during a single four-year term.
6. According to Moises Naim, Minister of Industry during the 1980s, policymakers’ complacency with oil-led growth-led policymakers to ignore the importance of coordinating or at least monitoring other indicators such as the exchange rate, interest rates, trade policy, or industrial policy. The result was rigid policy framework ill-suited to cope with external shocks such as the second oil crisis of the 1970s. See Naim (1993).
References


Cukierman, Alex. 1992 *Central bank strategy, Credibility and independence*. Boston: MIT.


Herrera, Genaro Arriagada, and Carol Graham. 1994. Chile: Sustaining adjustment during democratic transition. In Stephan Haggard and Steven


Asian Values or Asian Ideology? The Role of Culture in the International Relations of East Asia

Andrew Oros*

This essay explores the question of "Asian values" within the context of the changing international relations of post–Cold War East Asia. It begins by examining the advantages and drawbacks of using culture to understand international relations, arguing for the importance of a positivist methodology over a hermeneutic approach. It concludes that the idea of "Asian values" is based on the latter, and reflects more an ideology of certain Asian leaders than the cultural disposition of Asian peoples.

East Asia is at a crossroads today. The end of the Cold War has freed up options for many states in the region. Traditional global geopolitics no longer dictate state action. What this will lead to is a concern of policymakers around the globe. This realization that states have self-directed options rather than circumstances dictated by superpower rivalry has led some analysts away from the systemic-level approach of structural realism and down to levels of analysis that consider the role of culture, institutions, ideology, and individual beliefs. These aspects of international relations have been underdeveloped in the discipline, but are gaining new ground. Among them, the study of political culture is undergoing what one researcher has described as a renaissance (Inglehart 1988).

This essay focuses on one aspect of the evolving debate on the

*Ph.D. Candidate, Columbia University; M.Sc., London School of Economics; B.A., University of Southern California.
role of unit-level analysis in the international relations by examining the question of political culture in East Asia. Specifically, it centers on efforts to define East Asia as something other than simply a geographic concept. It will explore the question of "Asian values." First, however, one must consider the general challenges of using cultural over systemic explanations for regional international relations. For this purpose, the essay begins with a discussion of the current state of research regarding cultural approaches to international relations. Next, it considers the question of why cultural approaches are undergoing a renaissance today. Finally, it examines the debate over Asian values in light of the preceding analysis.

Cultural Approaches to International Relations

Cultural approaches to international relations have faced tough challenges in the field. Since the behavioralist revolution, theories in the discipline have been held up to a high standard derived from the scientific method. In addition, systemic theories of international relations have dominated during this period, largely due to the geopolitics imposed by Cold War bipolarity. Contemporary neorealist or structural realist theories are hostile to culture by the very nature of their systemic explanations. Culture has no place in an international system where interaction is determined by the relative power of the major states characterized as like units seeking greater power in the system (Waltz 1979; Gilpin 1981). Contending theories in the field have also tended to be based on the system level and are similarly hostile to culture. World systems theory and its cousin Dependencia theory, for example, hold that culture is irrelevant because the key factor in national development is the world capitalist system (Wallerstein 1974; Duvall et al. 1981).

Rational choice theory, which has been growing in popularity in the field over the past several decades, reinforces the belief among international relations theorists that cultural explanations have nothing to offer international relations theory. Rather than merely suffering marginalization by systemic theories, rationalist explanations of behavior have overtly challenged the belief that the substantive content of individual people's ideas—or their group manifestation in culture—are important factors to consider. To many economists, and to political scientists captivated by rational choice and related modes of thinking, ideas are unimportant or
epiphenominal either because agents can often correctly anticipate the actions of another (independent of cultural considerations) or because of some selective process which ensures that only agents who behave as if they were rational succeed (Goldstein and Keohane 1993). Today, however, a countertrend may be taking place in international relations. As rational choice attempts to answer bigger questions, its flaws are becoming more apparent. This has led some back to the concept of culture in an attempt to bridge the gap between the two seemingly disparate approaches.

In his review of recent cultural approaches to understanding China, Nathan begins by identifying the central problem facing the researcher attempting a cultural analysis: There can be no doubt that culture exists and that it affects behavior, but in what ways? He poses several key questions: “What do we mean by culture in the context of comparative statements? How can a culture’s distinctiveness be conceptualized? What is required to demonstrate that such distinctiveness exists, what it consists of, and what influence it has on the performance of societies?” (Nathan 1993, 923). These are tough questions which must be addressed if cultural approaches are to succeed.

Advocating a cultural approach to international relations is not to say that people do not behave in self-interested and broadly rational ways. Indeed, culture would not be an identifiable phenomenon if people within the culture did not behave in generally uniform ways given certain stimuli. Rather, a cultural approach to international relations takes issue with the idea that interests precede beliefs held by actors. A culturally informed approach to international relations would factor in the possibility of one’s beliefs or culture defining one’s interests. For example, Goldstein and Keohane write in their edited volume, Ideas and Foreign Policy, “Ideas influence policy when the principled or causal beliefs they embody provide road-maps that increase actor’s clarity about goals or ends-means relationships, when they affect outcomes of strategic situations in which there is no unique equilibrium, and when they become embedded in political institutions” (Goldstein and Keohane 1993, 3).

Similarly, when using culture—rather than individual beliefs—as an independent variable to explain action it is important to distinguish between the beliefs or ideas that motivate action and the behavior that this manifests. As Nathan points out in his review
of cultural approaches to understanding China, “When one wants to use culture as an explanation for behavior, one must define culture as a pattern of mental attitudes separate from the pattern of behaviors that such attitudes are thought to explain” (Nathan 1993, 924).

Nathan’s distinction between beliefs and values on the one hand, and behavior on the other is fundamental to the development of a scientific methodology to operationalize culture as an independent variable in international relations. Further, it allows him to distinguish between two types of cultural knowledge: hermeneutic knowledge and positivistic knowledge. Nathan describes hermeneutic knowledge as that derived from interpretation of a culture’s context and its inner structure. Positivistic knowledge, on the other hand, is based on empirical confirmation of hypotheses formulated in a way that allows for falsification (Nathan 1993, 924, 928).

Findings from extensive survey research conducted by Inglehart in the advanced European democracies lead him to assert that “examination of political culture is an essential supplement to the rational choice approach” (Inglehart 1990, 15–16). His book, Culture Shift, is a positivist rather than a hermeneutic study. It is supported by survey evidence from twenty-five nations, with substantial time series data from eight. Through his surveys, Inglehart finds significant cultural variations between the advanced industrial states, which he argues are closely linked to the viability of democratic institutions. Further, he sees the forces of modernity causing a “culture shift” to “post-materialist values” as the economic development of the countries studied continues. Most importantly for the political culture approach is his finding that differences between societies are not decreasing, as many theorists have suspected, but in many cases intensifying. His surveys indicate, for example, that “even in advanced industrial societies, religion not only outweighs social class as an influence on electoral behavior, but actually seems to be widening its lead” (Ibid., 15). His conclusion is that “different societies are characterized to very different degrees by a specific syndrome of political cultural attitudes; that these cultural differences are relatively enduring, but not immutable; and they can have major political consequences” (Ibid.). Inglehart’s findings are significant for two reasons. First, he claims to have shown real cultural differences between nation-states that
Asian Values or Asian Ideology?

Westerners have been intrigued by the "uniqueness" of Asian cultures since their discovery by European explorers. Historically, nearly all major examinations of Asian political behavior have taken as a given that Asia is fundamentally different from the West, including works by seminal Western political philosophers such as Marx (1853) and Weber (1919, 1920). Despite the growing popularity of rational choice theories discussed above, however, and the growing use of positivist methodology in studies of cultures, hermeneutic cultural explanations of Asian behavior predominate in international relations and comparative politics.  

A common error in cultural analyses in East Asia is to confuse hermeneutic and positivistic knowledge within the same study. In Nathan's words, "Knowledge about a single culture is sometimes restated as knowledge about two cultures, hermeneutic knowledge as positivistic knowledge" (Nathan 1993, 924). This is the mistake made by van Wolferen's deep problematizing of the Japanese state, compared to his singular notion of the West in his seminal study, _The Enigma of Japanese Power_. This is not to say that his findings are useless. Quite the contrary, van Wolferen's insights into Japanese society helped stir a revolution in American thinking about Japan. He successfully details how Japanese society is unique. One has to take as a matter of faith, however, his second claim that Japan is somehow "uniquely unique" (Van Wolferen 1989, 14). This is because empirically one cannot compare two things that are unique. Thus, a cross-cultural comparison begins with an idea of similarity, that a society's uniqueness comes from the different weight it places on values and beliefs that exist in all societies to some extent. As has been seen, however, few empirical cross-cultural studies between Asia and the West exist. Most knowledge of the region is based on the hermeneutic approach.

A RETURN TO CULTURAL EXPLANATIONS?

With the end of the Cold War, people around the world and particularly those in the United States are having great difficulty in understanding how radically different the world has become. One major difference is that all conflicts in the world are now genuinely regional and local, with the possible exception of those that involve
major raw materials (such as oil) or that have great ideological implications (such as those surrounding Israel, South Africa, or Political Islam) (Johnson 1995). As Betts summarizes, “The Soviet collapse makes the answers to some basic questions less obvious than they once seemed” (Betts 1993/94, 35).

There has been a proliferation of writings on the question of what type of international system is most likely to succeed bipolarity. Writers have taken a variety of perspectives, from idealism to hard-core realism, from a unipolar world to multipolarity. Especially in East Asia, a region growing in terms of relative power capabilities, the United States is being forced to justify its security and economic relationship with nearly every country in the region. Betts summarizes some of the key questions facing U.S. policymakers: “Is it now in the interest of the U.S. for China to succeed in economic liberalization and become more prosperous? or for Japan to become a normal state, developing a ratio of military to economic power comparable to that of other large, rich countries? or for Korea to unify? or for Taiwan to democratize, or for Vietnam to remain poor?” (Betts 1993/94, 35).

In U.S. circles, many are searching for a new enemy, even if this means the enemy within (i.e., American decline). Within the milieu of hegemonic stability theory and power politics, nationalism and religious fundamentalism, many have been drawn to the idea of a “clash of civilizations,” a phrase coined by Huntington (Huntington 1993a). “The Clash of Civilizations” sparked a debate still raging in *Foreign Affairs* regarding the relevance of culture to international relations in East Asia. Positions have ranged from Gibney’s assertion that “culture is not destiny” to Lee Kuan Yew’s response that “culture is destiny” (Gibney 1993; Zakaria 1994). Kim Dae Jung sums up the lack of consensus to date, asking in the title of his piece: “Is Culture Destiny?” (Kim 1994, 189). As with the bulk of cultural studies of Asia, all of these articles are based on the authors’ interpretation of values, events, and their significance—not on empirical studies of the role of culture in Asia.

**THE CLASH OF CIVILIZATIONS?**

“The Clash of Civilizations” has resonated not only in American society, but in policy-making circles around the world. As Huntington himself has noted in one of his many responses to critics of his argument: “Since my article was published in the Summer 1993
issue of *Foreign Affairs*, it has been the subject of scores and possibly hundreds of articles, symposia, and commentaries in countries throughout the world. It has been intelligently debated by presidents and prime ministers, scholars and journalists" (Huntington 1994/95, 177). More than any other debate in international relations today—except perhaps the larger framework of emerging nationalism, of which the clash of civilizations is itself a part—the idea of a clash of civilizations fills the need for a globally applicable, systemic-level explanation that could parsimoniously predict the course of international events for years to come. It harkened back to the good old days of predictable bipolarity.

The essence of Huntington's argument is not just that people are different but that groups of people are different—in uniform and identifiable ways. People have cultures and these cultures affect—indeed *define*—the way they act and interact in the world. His conclusion is clear:

It is my hypothesis that the fundamental source of conflict in this new world will not be primarily ideological or primarily economic. The great division among humankind and the dominating source of conflict will be cultural. . . . The clash of civilizations will dominate global politics. (Huntington 1993a, 22)

Thus, Huntington sees a grand progression in the history of the conflict among men from conflict between kings, to conflict between peoples, to conflicts between ideologies, to the apex—conflict between civilizations. Huntington describes a civilization as "a cultural entity" that is "the broadest level of identification" with which an individual can identify. He proposes seven or eight such civilizations: Western, Confucian, Japanese, Islamic, Hindu, Slavic-Orthodox, Latin American, and possibly African. The differences between these large blocs are fundamental, "differentiated from each other by history, language, culture, tradition and, most important, religion." In his view of history, "differences among civilizations have generated the most prolonged and most violent conflicts" (Huntington 1993a, 23–25). It is relevant to note at this point that many of these differences also apply *within* individual civilizations.

Huntington's formulation poses many problems. The key problem, addressed by this essay, is that Huntington is misguided in his attempt to find a new conflict cleavage to fit into the Cold War framework. The "clash of civilizations"—if indeed there is one—is
not nearly so overt. If there is any merit in his argument, it lies in the "soft power" level, which Huntington dismisses by writing that "the fundamental source of conflict in this new world will not be ideological." One region where a new ideological battle may be forming, however, is an area that Huntington sees as a "cultural threat": East Asia. In East Asia today one can hear a growing debate over "Asian values" and their superiority to Western conceptions of society. Thus, what one should be examining is not Chinese missile exports to other civilization groups (Huntington 1993b) but Chinese "values" exports to other states in the region. If there is a clash of civilizations in Asia today, it is an ideological clash, not a broadly cultural one.

What, however, is the competing ideology? During the Cold War, it was communism. All interstate conflict in the world was dichotomized into a communist/anti-communist struggle. The U.S. goal was "containment." Today, communism is dead. Freedom and democracy have prevailed. For a brief moment, some even contemplated "the end of history." It was this type of thinking, according to Mahbubani, that led to current problems, and ultimately perhaps to Huntington’s piece. Despite Huntington’s worries, Mahbubani points out, it is the West’s influence that is feared in most parts of the non-Western world (Mahbubani 1993, 10–14). Bartley’s critique of Huntington provides some examples of this. He writes:

The new Japanese crown princess was educated at Harvard, and the latest sumo sensation is known as Akebone [sic], but played basketball as Chad Rowen. The world’s language is English. Even the standard bearers of ‘the rest’ were largely educated in the West. (Bartley 1993, 16)

The West’s overpowering influence globally is especially evident if one equates modernity with the West. As Jeurgensmeyer argues in his variation of the clash of civilizations, the Western institution of secular nationalism has transformed the globe. The challenge between Western secular nationalism and traditional religious nationalism is difficult at best to accommodate and can never be fully reconciled (Juergensmeyer 1993). In the years since the end of the Cold War, various responses to this perceived overzealousness of the West can be seen in Asia. Quite the opposite of Huntington’s predicted unified voice, however, today one hears leaders from throughout the region asking, what is Asia? Can Asians be identified by common traits? Are there such things as ‘Asian
values? As Ajami rightly questions in his response to Huntington, “where is the Confucian world Huntington speaks of?” (Ajami 1993, 6).

THE SEARCH FOR ASIA

It is often repeated that the Asian challenge to the West is much more interesting than communism’s ever was because Asian authoritarians (unlike those of the former Soviet bloc) argue from a position of economic and social success. ("Competitive Order" 1992; Graybow 1994). Attempts to identify this challenge, however, have proven difficult. The crux of the problem of identifying a unified, Asian challenge is the tremendous diversity of the region. East Asia accounts for over a third of the world’s population. Forms of government in the region today range from liberal democracy to strict authoritarianism and absolute monarchy, with gradations of each form in between. Freedom House rankings of civil liberties and political freedoms find states in Asia at the full range of their seven point scale on both indicators. Moreover, states within East Asia cover the entire spectrum of economic development: displaying a thirty-fold per capita income gap between Asia’s richest and poorest states, from Japan’s per capita income of $30,000 to China and Indonesia’s per capita incomes of less than $1,000.

There is not much similarity between Asian states on which to formulate an Asian challenge. Still, many in Asia are attempting to develop an “Asian” identity. According to Japanese journalist Yoichi Funabashi, however, and contrary to the civilization paradigm of Huntington, Asia’s ‘Asianization’ is the result of the globalization of its economy and media, not from forces emanating from within Asia. In Funabashi’s view, “the region is not being ‘re-Asianized’: it is being Asianized” (Funabashi 1993, 79, 77).

Funabashi argues that central to Asian notions of development is that the economy should have priority—democracy and other forms of political and social evolution should be thought of in terms of how well they serve economic development. He has recently written about what he calls “a new Asian identity” and “a cohesive Asian worldview,” in which he would include South Asia. He describes the “Asian consciousness” as “animated by workaday pragmatism, the social awakening of a flourishing middle class and the moxie of technocrats, although still tinged perhaps by anticolonialist resentment, racism and indifference to civil liberties” (Funabashi 1993, 75). This view of Asians may differ from the view
Andrew Oros

of Americans hold of themselves, but hardly comprises fuel for a clash that will outweigh other factors.

One concrete example of an Asian cohesiveness for Funabashi involves security. He writes: "Most Asian nations are former colonies or protectorates, and their independence was gained through economic and societal development. They have a strong tendency to think of security not simply in military terms but as a synthesis of military, economic, technological and social strengths." Thus, "the new regional consciousness in Asia . . . stems from changes in traditional security attitudes." (Funabashi 1993, 80–81)

This constitutes more of a description of an impact felt on all actors in the region, than a cultural attribute. Beyond these very general characterizations, however, the idea of a common view of Asia or Asians (even East Asians) quickly disappears. Perhaps because of this difficulty, the focus in the debate today is not something that has been measured concretely. The emphasis has moved from a search for a common Asian identity to a search for "Asian values."

**Two Emerging Blocs?**

In writing about the first APEC leaders' summit in Seattle in November 1993, Graybow observed that "the remarkable diversity among the leaders underscored two powerful regional trends. The two largest blocs of leaders were a core of democrats and three authoritarian rulers: Indonesia's President Suharto, China's Jiang Jemin and Singapore's Goh Chok Tong" (Graybow 1994, 41).

As an indication of the relative distribution of power, he continued, it is interesting to note that the question of democracy was not on the agenda at Seattle.

The most vocal people to watch in Asia on the subject of Asian values are from the authoritarian camp, especially Lee Kuan Yew and Kishore Mahbubani of Singapore, and Datuk Seri Mahathir Mohamad and Datuk Seri Anwar Ibrahim of Malaysia. These two countries are in many ways competing for the right to articulate what is Asia. Lee, the former Prime Minister of Singapore from 1965 to 1990, is now Senior Minister. Mahbubani is permanent secretary in the Ministry of Foreign Affairs. Mahathir has been the Prime Minister of Malaysia since 1981. Anwar is the Finance Minister and was recently elected Deputy President of the ruling party and is therefore nominally in line to succeed Mahathir as Prime Minister.

Among pro-democrats, Kim Dae Jung, the former opposition
party leader in South Korea, is one of the most vocal on the world stage. The person perhaps most associated with the democracy movement in East Asia is Aung San Suu Kyi, who won the Nobel Peace Prize in 1991 for her unwillingness to recant her pro-democratic views in the face of severe government oppression. Unfortunately, she has been silenced through house arrest. Each of these leaders are eloquent and vocal speakers on the topic of Asian values, but each puts a carefully different spin on how the West should view Asia.

Mahathir argues that the real aim of Western protests over the environment and human rights practices in the region is to keep Asia down. He has called “arrogant” the belief that democracy is right for everyone. He is quoted in a recent interview as saying: “To us, democracy means the welfare of the majority. While the individual must have his rights, these must not extend to the point where they deprive the majority of their rights” (Crovitz 1994, 20). Freedom House surveys confirm this mixed commitment. According to its 1993–94 report, though Malaysia has allowed somewhat greater political freedom than Indonesia or China, “Prime Minister Mahathir has used the cover of consensual politics to stifle opposition.”

In recent years the government has cut back sharply on development funds to the two opposition-controlled states, Sabah and Kelantan, and has taken measures to limit debate in parliament, including reducing speaking time on the floor. (Graybow 1994, 44)

Anwar advocated a more measured view in a recent speech to the Institute for Policy Research for Development in Bangkok: “Properly instituted, democracy will ensure order and stability.” When speaking of the challenges ahead, however, he cited as “foremost of these is the creation and preservation of social order” (Anwar 1994, 34). Earlier in the year, he was more critical of so-called Western values on the economic front, saying “it is highly improbable that a purely free-market mechanism will fulfill the development objectives aspired to by developing nations in Asia.”

As the former leader of Abim, the Malaysia Islamic Youth Movement, Anwar has strong connections throughout the region and is seen as a regional leader of the next generation of political leaders in Southeast Asia.

Lee and Mahbubani both put forth a far more nuanced version of Asian values, laced with overt praise for Western achievements.
Their authoritarian overtones are in some ways not as strong as Mahathir’s and Anwar’s, but they still see democracy as a tool that can be adjusted to the necessities of social order rather than an ideal in itself. Lee tends to subscribe to Huntington’s notion of a Confucian civilization. In his recent interview with *Foreign Affairs*, he defines this group by saying “when I say East Asians, I mean Korea, Japan, China, Vietnam, as distinct from Southeast Asia, which is a mix between the Sinic and the Indian, though Indian culture also emphasizes similar values.” He includes in his “cultural backdrop the belief in thrift, hard work, filial piety and loyalty in the extended family, and, most of all, respect for scholarship and learning” (Zakaria 1994, 113, 114). Curiously, given his record in Singapore, he describes Asians as culturally disposed to limited government and faults Western liberalism for forcing government to play too large of a role in society.

Mahbubani articulates a more developed position in the *Washington Quarterly*, writing: “Many Western societies—including the United States—are doing some major things fundamentally wrong while a growing number of East Asian societies are doing the same things right. [The United States] needs to question its fundamental assumptions about its social and political arrangements and, in the process, learn a thing or two from East Asian societies” (Mahbubani 1994, 5). One may well ask, based on the above discussion, what exactly he expects the United States to learn from “East Asia.” In sum, he argues that the United States suffers from too much freedom and can learn methods of control and order from East Asia. “There can be no more vivid example of how a nation can become trapped by its ideology,” in Mahbubani’s opinion, than Americans clinging to the ideology of freedom (Mahbubani 1994, 9). Mahbubani, like Lee, is no fan of democracy, though he would not discard it entirely. “Consent of the governed is absolutely vital for any government,” he writes, yet “to have good government, you often need less, not more, democracy” (Mahbubani 1994, 16–17).

Kim has recently emerged as a vocal counterbalance to those challenging the basis of democratic rule. He contends that democracy is bound to spread through Asia as it has in Europe and the Americas. In response to the growing chorus supporting the claim of a different Asian political culture, one hostile toward democracy, Kim writes that contrary to Huntington and many Asian leaders (such as Mahathir and Lee), Asia has “a rich heritage of democracy-
oriented philosophies and traditions” (Kim 1994, 191). He finds similarity to the Lockean notion of sovereignty residing in the people in the Chinese philosophy of Meng-tzu two thousand years earlier. Moreover, he cites other Asian practices dating back millennia, such as the rule of law and the selection of officials through merit-based examinations, as supportive of the principles of modern day democracy. A full examination of these counter-claims is not within the scope of this essay, though it must be said that many of his interpretations of Chinese history run counter to accepted Western scholarship. However, the historical consensus about the merit of such claims is not necessarily as important as the idea that the democratic tradition could be traced back to such practices if leaders attempted to ground them as such.9

Kim’s opinion of Lee is clear: “Lee’s view of Asian cultures is not only insupportable but self-serving.” Later he adds, “The biggest obstacle is not [Asia’s] cultural heritage but the resistance of authoritarian rulers and their apologists” (Kim 1994, 190, 194). He points out the hypocrisy of Lee’s statements about the role of government in the West and East Asia,

Asian governments intrude much more than Western governments into the daily affairs of individuals and families. In Lee’s Singapore, the government stringently regulates individual’s actions—such as chewing bubble-gum, spitting, smoking, littering, and so on—to an Orwellian extreme of social engineering. In his native South Korea, he continues, each household is required to attend monthly neighborhood meetings to receive government directives and discuss local affairs. In Japan, he notes, government ministries play a significant role in regulating the domestic economy. And in China, of course, the government regulates almost all aspects of daily life from where you live, where you work, to how many children you are allowed to have. As Kim puts it, “such facts fly in the face of [Lee’s] assertion that East Asia’s governments are minimalist” (Kim 1994, 190).

Kim is not above proselytizing about Asian values, though. He asserts that “the proper way to cure the ills of industrial societies is not to impose the terror of a police state but to emphasize ethical education, give high regard to spiritual values, and promote high standards in culture and the arts.” He, of course, is speaking as the former opposition leader of an ethnically homogenous state. Even so, one wonders how “ethical education” and promoting “spiritual
values” will rest with the 25 percent of the Korean population who are Christian. Rather than joining the bandwagon of Western-bashing, though, Kim writes that “moral breakdown is attributable not to inherent shortcomings of Western cultures but to those of industrial societies; a similar phenomenon is now spreading through Asia’s newly industrializing societies” (Kim 1994, 191). Thus for Kim modernity, not culture, provides the explanation for social decay. The solution to the threats of modernity may lay in culture, however.

**CULTURE OR IDEOLOGY?**

Western critics of this new “Asian values” debate are increasingly viewing the rhetoric of Asian values as merely the result of economic self-confidence and a desire to fashion a response to Western (read U.S.) pressure on democracy and human rights. Indeed it is not hard to see that the movement by younger Asian leaders is an attempt to create an ideological framework to justify strong government (and hence their own political positions) in a more liberal economic environment. What is clear from the level of rhetoric about “Asian values” and the dichotomy of positions taken by Asia’s democratic versus its authoritarian leaders is that this debate has very little to do with culture, and very much to do with ideology. It is not a question of how things are, but how they should be.

The current Asian values debate is a political debate taking place in public between the leaders of both camps, not unlike the appeals made by the Communists to peoples in developing countries in the 1950s. Just as with the Communist challenge, there is no empirical evidence to support the claim that the people of Singapore, Malaysia, or China have values and beliefs that lead them to prefer to be oppressed. In fact, today’s authoritarians do not have the ability to control information to the degree authoritarians could in the 1950s. Examples of contrary “values” in these societies abound. Pro-democracy demonstrations in recent years in China, Malaysia, and Indonesia had to be violently suppressed by the government. In Singapore, the ruling party could only muster 58 percent of the popular vote in its most recent presidential elections (August 1993) despite the government-selected opposition candidate’s statement that the ruling party’s candidate was “far superior” (Freedom House 1994, 495). In both China and Singapore there is a widespread fear among elites of the moral decay of society. Internally, how to
develop the types of values that these leaders say exemplify the Asian way remains an enormous political question. This is compounded for leaders of the heterogeneous states of Southeast Asia, which paradoxically both Lee and Mahathir represent. In some ways, the idea of larger "Asian values" can be seen as an extension of programs designed to create shared domestic values.

Singapore has been at the forefront of countries in East Asia seeking to create a "national ethic," initiated by then Prime Minister Lee. The goal in Singapore is to help Singaporeans keep their Asian bearings as they enter the 21st century: to find a way for Singaporeans to modernize, but not Westernize. Lee has blamed the use of English as part of the general problem of too many Western influences on Singaporeans. He has publicly lamented that the consequence of the widespread use of English among Singaporeans has been to give every citizen "a pocket translator" to explore the Western world. To combat this trend, Lee initiated a "Speak Mandarin" campaign in the mid-1980s and later introduced Confucian philosophy as a core subject in secondary schools. The government is not apologetic about its common censorship of "sensitive topics" and is proud that it stops (or persuades to stop) publication of anything that might stir up ethnic discord; ironically, though, in this area it is the government itself that is stirring up such discord through its "national identity" program which Malays and Indians say stresses Chinese values too much (Pan 1989; Balakrishnan 1989a, 1989b).

Similarly, in China there is widespread concern among elites that market reforms are creating Chinese who value only money. "Chinese people are now moving to a period of pragmatism and self-interest," remarked one well-positioned scholar in the Chinese Academy of Social Sciences. He noted that when the Institute for American Studies began publishing a new magazine for popular consumption called *American Panorama*, the topics of greatest interest to its readers were guns, crime, and sports (Wang Jisi, interview, 23 June 1994). The prevalence of such a modernized, counterculture—counter to Asian values—is not isolated to these two countries. The exact level of support for each set of views, however, is unknown. This is the drawback of such hermeneutic evidence. Until the debate on Asian values moves to a research program on Asian values, it is impossible to know the actual state of affairs.
**CONCLUSION**

Cultural explanations for states’ international behavior have been on the increase in recent years, primarily as a result of the end of the Cold War and the concomitant change in geopolitics. In looking at the policy ramifications for the major players in East Asia, however, the cultural approach as used to analyze recent rhetoric merely reflects cleavages already apparent due to existing political and economic rivalries and lends little insight into the international relations of the region. Cultural explanations based on an “Asian identity” are premature and reflect more the aspirations of (especially authoritarian) Asian leaders than the cultural predispositions of the people of Asia. In order to understand the cultural inclinations of Asians, one must propose falsifiable hypotheses, not derive a definition of culture through observation of behavior or blind adoption of any one person’s views.

The debate over Asian values is important to keep in mind, however, because it may grow into an ideology that exerts influence over the region. As the military and economic power of East Asian states grow, the spread of an ideology counter to American views represents a potential threat that should be monitored. Such an ideology, coupled with a potential hegemon such as China, would be a dangerous combination. This is not the situation today. The international relations of East Asia should not, therefore, be characterized as a “clash of civilizations.”

**Notes**

1. Andrew Nathan distinguishes between the “diminished positivism” more common in theory today and the natural science-like laws popular into the mid-1960s. Both, however, are equally antagonistic to hermeneutic theories of culture that have comprised the bulk of the literature on culture in international relations. Nathan describes “diminished positivism” as that which “requires only that any proposition be precisely specified (for example, with respect to time frame, social actors, geographical location, and indicators of amount or degree), that it be stated in a form that is potentially disprovable by reference to empirical evidence (the falsifiability criterion), and that it be treated as unproven until empirically proven” (Nathan 1993, 928).

2. For a recent example of this approach in the Asian context, see Ramseyer and Rosenbluth (1993), in which a rational choice, interest-based approach is used to explain Japanese politics. This leads the authors to the conclusion that Japan’s bureaucrats (conventionally believed to be powerful
Asian Values or Asian Ideology?

players) are merely the servants of LDP politicians.

Admittedly, this is a positivistic approach to culture for it assumes the ability to separate between beliefs and interests. The reflectivist school would reject this separation, and concentrate instead on how preferences are formed and how identities are shaped. For an introduction to this latter approach, see Wendt (1987).

Most of these studies focus on only one country in Asia. For one attempt at pan-Asian hermeneutic theorizing, see Pye (1985). Survey research in Japan is one obvious exception to this general rule. Many studies of Japanese voting behavior have been conducted, as have surveys measuring values and beliefs more generally. Japan has even been included among a growing number of countries surveyed under The Civic Culture framework. Additionally, in recent years survey research has been conducted to a limited degree in China; see Nathan and Shi (1993). There have been no pan-Asian, positivistic studies of culture similar to Inglehart’s work on Europe.

Joseph Nye defines “soft power” as “the ability to establish preferences,” often associated with “intangible power resources such as culture, ideology, and institutions” (Nye 1990, 32).

Freedom House surveys annually rank both “political rights” and “civil liberties” on a seven-point scale. From these rankings, they determine whether a country is “free,” “partially free,” or “not free.” Of the major states in East Asia, only Japan and South Korea are designated “free.” All other states are classified as “partially free” with the exception of China and North Korea which are “not free” by Freedom House standards.

Prime Minister Mahathir of Malaysia refused to attend.

This was the view put forth at a conference held earlier this year in Kuala Lumpur for Asians only entitled “Asia in the 21st Century.” See the February 10 issue of Far Eastern Economic Review for more details.

This argument is based on the idea that all national identities are “constructed” by elites to serve a particular purpose, usually to legitimate a claim to power. What is important is not so much the “truth” in the claim (there can be none) but that the elites are able to ground their claims in some aspect of previous history of the group (Anderson 1983).

References


Asian Values or Asian Ideology?


Traditional approaches to the question of how urban governments respond to fiscal stress have emphasized either the exogenous economic determinants of a city's fiscal environment, or the constraints imposed on elected officials by the demands of maintaining successful electoral or governing coalitions. Some writers have sought to define a space in which local governments can make their own choices, at least partially independent of economic or political constraints. Their efforts are not convincing, in part, because they ignore the additional constraint of the U.S. federalist structure. This paper asserts that the U.S. federalist structure plays a major part in the ongoing fiscal strains that cities continue to face in the 1990s.

This review essay explores the response of cities to fiscal stress. Numerous explanations have been advanced concerning the causes of fiscal stress, from public choice-type explanations to socioeconomic structural models. The question here concerns how city governments, confronted with rising demands for services and relatively declining resources, have responded. Most cities have adopted a "corporate-center" (Reed 1984) or economic development strategy, focusing on promoting growth of the central business area as the key to raising revenues and preserving services. In what ways does the organization of political power in a city affect...
how it has managed and responded to changes in its fiscal condition? Is there room for a different strategy, a different coalition of groups, to address the fiscal and social problems of cities?

The financial and other problems of cities have received extensive attention since the late 1960s, as a constellation of external factors began to change the demographic, social, and economic makeup of the nation's large urban centers. Federal policies accelerated these trends, and federal policies were put into place to address some of the problems they created (Barro 1978; Levine and Williams 1992). These extensive changes—both the larger socioeconomic trends and the policy responses from Washington—shaped the political environment in urban centers in new ways. The political challenge to city-governing regimes was to negotiate these transitions as smoothly as possible. Some cities were more successful than others at doing so. The failure to achieve a successful transition was sometimes manifested by civil unrest. More often, however, it took the form of a chronic state of fiscal stress, in which the municipal government consistently faced service demands that exceed its ability to provide them out of the resources available. Occasionally, the fiscal stress became so acute as to erupt into a crisis and the risk of default.

As more and more of the nation's largest cities began to experience chronic or acute fiscal problems in the late 1960s and early 1970s, numerous studies sought to identify the causes of this phenomenon. This paper will begin by reviewing some of the explanations for the causes of fiscal stress. Then, a brief review of the current fiscal condition of large cities will be undertaken. This review indicates that the decade of prosperity of the 1980s has not alleviated the underlying imbalance in urban fiscal conditions.

Next, the paper will examine alternative models of response to fiscal stress. These fall into two broad categories: organizational and administrative explanations, which focus on city governments as administrative and bureaucratic organizations, and political science explanations, which seek to explain the response to fiscal stress based on the city's political organization and sources of political power. The last section will utilize regime theory to explain how urban governments manage fiscal stress. If "government authority is inadequate for governing," as regime theory states, and the "cooperation and participation of nongovernmental [private sector] actors becomes essential," what form does this cooperation
and participation take with respect to making choices in the face of fiscal stress and the imbalance between needs and available resources? Is it possible to have a regime that includes members other than downtown business and financial interests and government? Is it possible to have a progressive regime? Or are cities condemned inevitably to reliance on a "corporate-center" strategy? This paper argues that, under the current structure of American federalism, the imperative of competing for private investment and growth is virtually unavoidable and leaves little margin for other municipal priorities.

**Urban Fiscal Stress: Causes and Consequences**

The local government sector grew more rapidly than either the state or federal government sectors during the 1960s and early 1970s, especially in the nation's largest cities (Ladd 1994, 203). Urban fiscal stress was nearly a constant throughout the same period. Some cities have coped better with the squeeze on urban financial resources than others, with New York's 1975 bankruptcy providing only the most dramatic example of the latter. Five types of explanations are usually offered for the growth in city spending and the fiscal strains that seemed to accompany it during this period. These are: socioeconomic—changing mix in both population and economic activity; rising public sector costs; the impact of federal assistance on city budgetary choices; the public-choice school model of bureaucratic growth; and finally the "political vulnerability" of city hall (Rubin 1982).

The socioeconomic explanation focuses on the migration of poor blacks and Hispanics to big cities in the Northeast, Midwest, and far West in search of jobs, and on the simultaneous out-migration of the middle and upper-middle classes to the suburbs, spurred by federal-highway and mortgage-insurance programs (Vaughan and Vogel 1978; Rubin 1982). The result, a changing urban population mix, put pressure on both the expenditure side and the revenue side. Demands for services rose as the population became poorer. Simultaneously, declines in property values, average incomes, and retail sales all reduced the tax base. As a consequence, available revenues to meet these demands fell. In order to raise revenues, tax rates were increased, further propelling the middle class and businesses out of the city, in a kind of negative feedback loop.
Another explanation focuses on the rising costs of delivering city services. This is principally a result of the growing power of public-employee unions and their importance in city politics. Negotiating substantial pay raises and restrictive work rules pushed costs up without necessarily increasing productivity.\(^1\)

Generous federal aid to cities for poverty and urban renewal programs in the 1960s and 1970s also helped fuel an expansion of services, especially to the poor. This increase in demand helped raise wage levels in the public sector and created new service constituencies. The reduction of federal and state aid in the 1980s and 1990s left cities with considerable service burdens without the resources to finance them. As one observer put it, “When federal aid began to level off in the early 1970s, the claims it had generated did not follow suit” (Pecorella 1984, 307–8). In addition, the taxpayer revolt of the late 1970s, resulting in various statewide spending or tax limitations, further constrained the ability of cities to make up the lost revenues with general purpose taxes.\(^2\)

The public-choice school of nonmarket economics has put forward its own explanations based on bureaucratic growth. Two variations of this theory emerge. First, citizens demand too high a level of services, or too many services, because they don’t bear the price directly. This is especially true of services that benefit a subset of the population but are paid for by everyone. Second, self-interested bureaucrats and public officials, who have no other measure of success than the size of their own budget, seek to maximize their department’s total share of the annual budget.

The political vulnerability of city hall to demands for services is another explanation for rising expenditure levels: “[T]he more vulnerable city hall is to demands the greater the likelihood of fiscal stress” (Rubin 1982, 8). One variation of this concept is that growth in spending is explained by “waxing and waning coalitions of voters.” Mayors increase spending in order to obtain the votes of various groups. In the 1960s, this was exacerbated by the increased participation rates of minorities (Reed 1984). Vulnerability may also be explained as a function of the form of government. According to this explanation, “machine and strong-mayor governments are more responsive to, yet more vulnerable to, demands than reformed, city-manager governments” (Rubin 1982, 9). Finally, the relative strength of competing interest groups determines who gets what. The more organized and active groups are better at getting what they want from city hall.
Recenf Trends in the Fiscal Condition of Cities

Although the first signs of fiscal stress appeared in the late 1960s and early 1970s, it was the New York City fiscal crisis of 1975 that first drew acute scholarly attention. In the twenty years since the New York City fiscal crisis and the beginning of the decline in federal aid to cities, spending growth in big cities slowed considerably. Ladd (1994, 216) notes that “the growth in big-city spending lagged behind that of total spending in a sector [state and local government] that was declining relative to the rest of the economy.” As the economy performed well during the 1980s, big cities saw something of a resurgence in their spending, although at a slower rate than the combined state and local sector or the local government sector taken as a whole. The following table documents these trends.

**Growth in Big City Spending Relative to State and Local Sectors, 1977–1989**

<table>
<thead>
<tr>
<th></th>
<th>General Spending Per Capita (1989 dollars)</th>
<th>Growth Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All state and local gov’ts</td>
<td>2,569</td>
<td>2,611</td>
</tr>
<tr>
<td>All local governments</td>
<td>1,633</td>
<td>1,566</td>
</tr>
<tr>
<td>32 big cities*</td>
<td>1,106</td>
<td>1,073</td>
</tr>
<tr>
<td>New York City</td>
<td>3,413</td>
<td>3,165</td>
</tr>
</tbody>
</table>

*Excluding New York.

Despite a growing economy and moderate increases in spending, Ladd’s analysis of the financial condition and underlying fiscal health of big cities leads her to conclude that the financial condition of big cities was “not particularly strong” at the beginning of the 1990–91 recession. Their underlying fiscal health—their ability to raise sufficient revenues to meet expenditure needs on a standardized basis—was also poor. Indeed, the larger the city size, the worse off it appeared to be. The fiscal conditions of cities with over a million population were considerably worse than small and medium-sized cities. Drennan (1984) and Brecher, Horton, and Mead (1984) document this disparity for the city and state of New York, noting that while the economic growth of the 1980s reduced
deficits, the downturn of the early 1990s once again revealed the long-term structural imbalance between revenues and expenditures. The overall fiscal problem of cities has not been solved, and they remain quite vulnerable to cyclical downturns as well as long-term structural economic changes. This persistent imbalance between revenues and expenditures has promoted a considerable body of literature, which attempts to discover and explain the nature of cities’ response to fiscal stress.

ALTERNATIVE MODELS OF RESPONSE
Not surprisingly, the least fiscally healthy cities on Ladd’s fiscal health index imposed the highest tax burdens on their residents: “The data suggest that big cities apparently have responded to their declining fiscal health both by raising tax rates and reducing service levels.” (Ladd 1994, 254). This general statement conceals the variety of possible measures cities can take to respond to fiscal retrenchment. Pammer (1990) details a variety of strategies that cities employ for coping with fiscal retrenchment. These fall into three broad categories: (1) revenue strategies, such as raising taxes or diversifying revenue sources (including obtaining more intergovernmental aid); (2) productivity improvements and alternative service delivery measures (e.g., privatization); and (3) expenditure cuts, i.e., cutting services, either in an across-the-board fashion, or with more targeted program cuts. These various measures can be employed in any combination—suggesting, as Pammer notes, that the choice of strategies is governed by something other than purely economic criteria. The question of why cities choose particular strategies and reject others has been the subject of considerable research.

Organization Theory Analyses
One group of analysts draws on organization theory to explain the response of city governments to fiscal strain, and to explore strategies of retrenchment. Levine et al. (1981, 16), for example, “rely heavily on the conception of a local government as a type of complex formal organization that shares properties and behavioral patterns with other organizations and can be expected to adapt to changes in resource levels in predictable ways.”

Curry (1990) uses a “political or organizational anthropology” to study four California school districts in the wake of Proposition 13.
His thesis stands in contrast, he says, to the standard expectations of organization theory and of "recent examinations of organizational responses to fiscal stress," which state that organizations will try to preserve their "basic structure for substantive goal attainment." He concludes that problems of fiscal stress will divert attention away from basic organizational goals: "When confronted with fiscal stress, conflict management may be of primary concern, not subordinated to substantive achievement." The conflict arises from the disappointed expectations of members of the "organizational coalition." In times of plenty, this manifests as decentralization and diffused conflict; in times of scarcity, by contrast, decision making is centralized and conflict among the interested parties increases.

Pammer further focuses on administrative processes and explicitly adds political variables. He summarizes Stonecash and McAfee's argument that "while socioeconomic forces may contribute to fiscal problems, these factors represent conditions to which city political systems may or may not respond . . . [T]he reality is that cities respond differently to changes in their resource levels . . . variations in responses are attributable to the political decisionmaking processes characteristic of cities" (Pammer 1990, 9–10). Additional factors he hypothesizes include socioeconomic characteristics of the city (older, larger, Northeast-Midwest manufacturing cities in a more acute state of fiscal strain) and the role of interest groups (ethnic or neighborhood groups, business organizations, public-employee unions, department administrators, advocacy groups, taxpayer/homeowners groups, etc). Nonetheless, he concludes that "administrative variables exhibited a stronger effect on retrenchment actions than the environmental decline measure. To some extent this finding supports the notion that local officials play a more important role in the use of strategies than socioeconomic factors" (Ibid., 46). Moreover, "political factors do not play an important role in dictating retrenchment policies; instead the chief executive will be the dominant figure in the process, so whatever action is taken will reflect that person's perceptions of what the situation demands" (Ibid.). Levine et al. found a similar trend toward centralization of authority during the period of retrenchment and a decline in the role of interest groups, which the authors attribute to a narrowing of the scope of government and the transfer of governmental functions to other levels (Levine et al. 1981, 203).
Most authorities on the budgetary process have concluded that increased centralization of decisionmaking will result from fiscal stress: "Scarcity tends to impose more centralization in allocative decisionmaking" (Rubin 1993, 26). In contrast, Caiden (1980) found that formal governmental authority can become less effective during periods of moderate fiscal stress as interest groups exert pressure on public officials. The New York City experience indicates the most extreme version of loss of political authority. The mayor was virtually stripped of budgetary power in favor of the Emergency Financial Control Board and the Municipal Assistance Corporation—structures dominated by banking and business interests and imposed on the city by Albany to approve city spending plans.

Clark and Ferguson's (1983) more complex model takes what they call a "systems analysis" approach to understanding the dynamics of fiscal strain. In their model, citizen preferences are expressed through organized group pressure on political leaders, who more or less faithfully reflect these preferences, depending on how dynamic they are in imparting their own. The fiscal decision "outputs"—tax rates, expenditure allocations, debt levels, etc.—then feed back into citizen preferences, resulting in and combining with migrations into and out of the city, and the reformulation of preferences among sectors of the population (e.g., the middle class, minorities, municipal employees). Clark and Ferguson's view of government is "a system which must adapt to its environment"; their conception of fiscal strain "is the degree to which government expenditures and debt are adapted to the city's private sector resources." This adaptation is conceptualized as a series of ratios between fiscal outputs (general expenditures, long-term debt, own-source revenues, and common functions) and private sector resources. The private and public sectors are distinct and thus their ratios change through independent processes—the former "political," the latter minimally so. The processes and interactions among the components of this system are defined by "transformation rules" characteristic of four political cultures: New Deal Democrats, Ethnic Politicians, New Deal Republicans, and New Fiscal Populists.

**Political Science Approaches**

The notion of transformation rules and political cultures forms a more traditional political science approach. Political science stud-
ies tend to focus on the influence on decision making of nongovernmental actors and their interactions with elected officials. How organized interests interact with elected officials in the political process to determine outcomes can be influenced by several factors. The structure of municipal government is found by some to be an important determinant of response. So-called reformed governments are supposedly less vulnerable to political pressure than unreformed machine or patronage-based governments. “Research . . . has shown that formal structures of cities advantage some interests and disadvantage others in the policy process which ultimately affects the kinds of fiscal policies cities adopt” (Pammer 1990, 48). Reformed governments are more centralized, more efficient, less subject to political pressures, and thus should respond better to fiscal stress.

Results [of research by Lineberry and Fowler] showed that cities with reformed institutions tended to be less sensitive to the higher spending demands of lower-income and minority interests. Hence, contrary to the public-regarding view, reform cities tended to impose less tax and expenditure burdens . . . The finding that reformed cities produce lower spending commitments suggests that perhaps these cities are better suited than unreformed cities to manage fiscal stress. Reform governments typically embrace the idea of efficiency, sound budgeting practices, and centralized control over the budget and personnel. (Pammer 1990, 49)

In general, these attributes should allow for more effective cutback management with less interest-group influence on the outcomes.

Some researchers have focused on the budgeting process as a political process and have found that, whatever political culture or organizational structure prevails, attempts to respond to fiscal strain will usually proceed incrementally, rather than through any dramatic restructuring of revenue sources or expenditure patterns. It follows, then, that the best predictor of this year’s expenditures is last year’s (Pammer 1990; Rubin 1980; Schick 1980). This runs somewhat counter to what two observers describe as the “long and multidisciplinary tradition of the [over]responsive local public sector” (McDonald and Ward 1984, 29–30). This very stability in budgeting might help explain, however, why cities get into a condition of fiscal stress: policy does not change much even in the face of evidence that expenditures and revenues are increasingly out of balance. Incrementalism leaves open the question of why
public officials behave incrementally. Is it because of the cautious and self-protective behavior of bureaucrats? Or because of the need to avoid upsetting established electoral coalitions? Rubin (1980) argues that officials are more inclined to "muddle through," rather than make the politically unpopular choices they face in a period of fiscal stress. McDonald and Ward find a confluence of factors that explains this behavior: "For better or for worse, it was restrictive fiscal ideologies, vested bureaucratic interests, and failure to achieve consensus on fiscal expansion that made local politicians timid, seemingly inert, and always incremental" (McDonald and Ward 1984, 32).

Another factor cited to explain the success or failure of urban politicians in responding to fiscal strain lies in the scope of city responsibilities. Cities with a narrow range of functions may have a more manageable governance problem than cities, like New York, with a very broad range of responsibilities. If cities are burdened with redistributive programs, then they will feel the strains of managing these tensions more acutely (e.g., New York in 1975). This is not least because many social service functions are joint responsibilities of the city and with state and federal governments, greatly restricting the city's freedom to control its own budget allocations. If they are not so burdened, then the fiscal situation might be relatively better (Detroit, with a narrow functional scope, is perhaps one exception).

Some studies have found that differences in functional scope are related to different political cultures associated with geographic regions. Older Frostbelt cities of the Northeast and Midwest are characterized as disposed toward high levels of social service provision, an acceptance (however begrudging) of the mobilization of both economic and racial minorities, a recognition of the legitimacy of municipal employee unions, a tolerance for local patronage systems, a suspicion of the motives of businessmen, and an indifference to the maintenance needs of the city's capital infrastructure. (Sbragia 1983a, 2)

This stands in contrast to the newer cities of the Sunbelt in the South and Southwest. The Sunbelt cities are "attractive for sunshine, low taxes, and a pro-business climate. . . . More so than the Frostbelt, the South and West are suspicious of government intervention to help minorities and the poor, hostile to taxation, pro-business, reluctant to accept public (or private) sector unionization,
and uncomfortable with collective as opposed to individual solutions to problems" (Ibid., 3). Sbragia points out that the Sunbelt cities also have generally enjoyed the advantage of being able to annex growing areas, and thus avoided the central-city/suburb disparity that plagues many Northeastern cities.⁴

Echoing Sbragia's characterization, Shefter writes:

American cities fall into one of two categories. The first, which is most typical of the Southwest, is composed of cities in which local business elites dominate local politics, and their priorities are reflected in municipal fiscal policy. The second category is composed of cities, such as New York and other major cities in the Northeast, in which the local political system possesses a significant measure of autonomy from the hierarchies of civil society and in which politicians and public officials are at times tempted to engage in deficit financing. In such cities fiscal discipline is imposed upon the local government by the municipal bond-rating agencies and the public capital market, and by the threat that if the market closes to the city, political forces allied with local business elites may be able to wage a successful reform campaign and gain control of City Hall. Such threats provide the city's politicians and public officials with a strong incentive to reach an accord with business and to pursue fiscal policies that are acceptable both to it and to participants in the public capital market. In other words, municipal governments may be subject either to internally-imposed fiscal discipline or to externally-imposed fiscal discipline, but in one or the other form they cannot evade such discipline over the long run. (1985, 232)

**Regime Theory and Response to Fiscal Stress**

Political scientists, then, have usually focused on governmental structures and on interest groups as actors in the formation of policy, including fiscal policy. Rubin (1993), for example, characterizes budgeting as a process involving multiple actors, with numerous and varied motivations, and differential power over time. Different actors are involved in different parts of the process: revenues, expenditures, policy, implementation, budgetary balance. The power of interest groups is seen as principally electorally based. There is a certain determinism to this approach, as noted by McDonald and Ward:

Most historians of the politics of urban fiscal policy assume that once one has identified the social basis of support for an urban regime, one
has also identified the nature of the regime’s fiscal policy. This follows because of the essentially functionalist notion that political and social institutions persist because they fulfill certain functions for their supporters. The implicit slogan of functionalist analysis is, ‘Survival implies ongoing usefulness—search it out.’ However, because of the variety of variables intervening between them, there is no necessary connection between a political movement’s social basis and the nature of its policy. To believe otherwise is . . . to assume that political institutions mechanistically reflect the social configurations that underlie them. (McDonald and Ward 1984, 2)

The literature on responding to fiscal stress that focuses on political variables tends to see nongovernmental actors as constraining. Unions, minority groups, business groups, and other political claimants are modeled as if the government sector had some objective function to maximize.

In contrast, regime theory in Stone’s formulation (1986, 1993) is based on the notion that “government authority is inadequate for governing”; and therefore the “cooperation and participation of nongovernmental actors becomes essential.” In Stone’s conception, the coming together of different groups empowers both government and its private sector allies, rather than constraining them. He calls this process a social production concept of power, as opposed to the traditional social control notion. The power to govern is not “captured” but “created by bringing cooperating actors together, often as unequal contributors to a shared set of purposes” (Stone 1993, 8). In particular, businesses and the resources they control are too significant to be excluded from the job of governance; they enjoy a “systemic power advantage” not commensurate with the small number of votes they actually control (Stone 1980). Significantly, the electoral coalition and the governing coalition therefore need not be—and indeed probably will not be—one and the same.

The social production model helps explain why regimes can endure difficult environmental changes and stresses. However, this view of how urban regimes operate, maintain, and adapt, was developed from looking at a city where fiscal stress was not a preeminent problem. Much of the durability of particular coalitions between business and government in the economic development sphere is premised on the creation of particularistic benefits: “selective incentives” and “small opportunities” as Stone calls them.
In regime theory, “politics is about the production rather than the distribution of benefits” (Stone 1993, 8). In a city experiencing fiscal stress, the challenge of “producing benefits” is intensified. In a situation of fiscal stress, the availability of these benefits may be limited and perhaps diminishing. The city is, however, confronted with the dilemma of meeting service demands, which can pit the needs of maintaining a coalition against the potentially competing demands of its members, in a context of diminishing resources. If a regime is based on more than a mere exchange of votes for services, for example active cooperation of the participants toward shared purposes, can these shared purposes continue to be pursued when the resources of government are highly limited?

Stone notes what he calls the “two basic institutional principles of the American political economy: (1) popular control of the formal machinery of government; and (2) private ownership of business enterprise” (Stone 1989, 6). It is the tension between maintaining the city’s economic viability, on the one hand, and meeting the demands of its population for services and political participation, on the other, that characterizes the American city’s governance problem.

Echoing this theme, and borrowing from the neo-Marxian theory and arguments of James O’Connor (1973), Pecorella (1984) argues that the problem of governance is in balancing the demands of two broad constituencies: business and financial interests, who enjoy the advantages of “systemic power” (he borrows Stone’s term) through their control of the capital that makes them crucial to economic development, on the one hand, and “attentive non-elites,” on the other—the “politically organized elements among the urban masses” whose power is principally the ability to generate votes. These groups come into conflict during retrenchment, as the latter seek to protect their gains, and the former to ensure economic and political stability. The Pecorella model of retrenchment politics “weighs the import of business demands for cutbacks and government reorganization as more significant and direct than the counterclaims of attentive non-elites for service maintenance and participation . . . [but] does not treat decisionmakers as ‘tools of elite interests’” (Pecorella 1984, 299-301). Friedland, Piven, and Alford share this view:

The dilemma of cities in capitalist societies is how to maintain a structure of expenditure and taxation that can stimulate stable eco-
nomic growth, while at the same time maintaining the popular legitimacy of governmental institutions, even when the potential for political conflict becomes intense. The series of structural arrangements [they describe] tend to convert political conflict between groups and classes into demands on the state which force state expansion. But this process also tends to create fiscal strains. Public expenditures increase faster than the state's ability to finance them from its own revenues. Thus fiscal strains are a recurrent feature of capitalist cities. (1978, 219)

Cities must balance the imperatives of the credit markets and economic development against the demands of citizens, expressed through organized groups acting by various means on the formal structures of power for greater levels of services. The resources that businesses control are crucial to the city's economic well-being; the votes that organized groups of citizens control are crucial to reelection.

This raises the question of just how much freedom of maneuver is available to politicians, and whether they can effectively control the municipal agenda. Peterson (1981) and others have suggested that city political leaders are virtually powerless in the face of economic "restructuring" and other phenomena that are exogenous to the local economy. They seek to define a space for politics in the general urban crisis. Numerous authors have responded to this assertion. In his case study of Cleveland's 1979 default, Swanstrom (1985; 1986) criticizes the recent trend toward economic explanations of urban politics as creating "an abstract economic determinism," which he summarizes thus: "Local policymaking is tightly constrained by the functional need to serve mobile wealth . . . Policies are not viewed as the result of the intentions of political actors but as the product of the function of the political system in the economy" (Swanstrom 1986, 81–82). Whether Swanstrom's critique is neo-Marxian or follows Peterson's analysis, they both assert that the modern city is highly constrained in its policymaking freedom by the imperatives of economics.

To the contrary, Swanstrom asserts that Cleveland's default was largely political. Motivated by the animosity of a tightly knit business community toward a populist mayor, the community used their financial power to refuse to roll over city notes unless certain conditions (the sale of the municipal power company to private interests) were met. The banks and other financial interests made
a political choice that threw Cleveland into default; they could have chosen not to call in their loans, as had been done in other cases. “In short, the economic pressures on the political system were mediated in crucial ways by politics . . . [T]he degree of discretion that is available to local governments and to investors must be determined empirically” (Swanstrom 1986, 102–3). Economic pressures are not unimportant for political decision making, but their effects are mediated in complex ways by political variables.

Stone also seeks to refute an economic determinist argument by describing the workings of a political regime consisting of an alliance between elected officials and downtown business interests. But while he tries to explore an alternative to pluralist and elitist theories of urban governance, and to show the empowering nature of a coalition regime, he also demonstrates just how constrained politicians are by the imperatives of economic growth, and the systemic power advantages of business. His chapter on Maynard Jackson’s election to the Atlanta mayoralty is particularly revealing. It describes the business-government alliance that reasserted itself in the face of an attempt to set the political agenda by Jackson’s electoral coalition.

Federico Peña was in some respects the Maynard Jackson of Denver. He came into office with the support of a diverse coalition and “promised to open city hall to neighborhood groups, minorities, and others who had previously been shut out of city government” (Judd 1986, 152). Like Atlanta, Denver had a history of pro-growth, downtown development policies, with neighborhood groups and minority groups largely shut out of city policymaking. Peña initiated neighborhood “town meetings” and increased the size of the neighborhood planning staff. However, before the end of Peña’s first term, economic development had become the administration’s first priority. The city was compelled by the logic of a declining revenue base. Investment to make the city an attractive location for both business and more affluent residents was necessary to maintain spending levels—and to avoid the vicious circle of declining services leading to residential and business flight, further undermining the revenues to provide services. The Denver equivalent to Atlanta’s CAP—the Denver Partnership—and the mayor, formed a commission to create a master plan for development of the central business district. Denver Partnership members outnumbered neighborhood representatives on the
steering committee. The Denver Partnership provided 75 percent of the initial funding and two-thirds of the staff for the planning project. The city's chief planner explained that the partnership had a dominant voice in the downtown plan because the city did not have the core staff or the money to do the work on its own (Judd 1986, 155). Again, the "systemic power" advantage of business in money, personnel, and other resources is illustrated. Judd attributes the failure of neighborhood and minority groups to mobilize against this policy thrust to the mayor's ability to convince voters that downtown development was necessary before the city could afford to improve or increase neighborhood services. As an aide to the mayor put it, "We probably will be criticized by some for focusing too much on downtown. People will ask, 'What does Saks Fifth Avenue have to do with the neighborhoods?' The answer is that downtown retailing has everything to do with the neighborhoods. The city needs the revenue" (Judd 1986, 145).

Cities are thus caught in the tension between the need to promote themselves to the private sector—the development strategy—and the need to manage conflict and promote social cohesion. The latter is particularly true for mayors elected by a voter coalition consisting of those who do not directly benefit much from pro-development strategies, who "may be impatient for the benefits of growth to trickle down" (Ibid., 148).

Shefter (1985) sees the resolution of this tension, at least in the case of New York City, in a cycle between reform and post-reform regimes. No matter the party label or rhetoric of successive candidates for mayor, the governing regimes they formed reflected their position in the cycle between periods of fiscal plenty and periods of fiscal austerity.

Shefter finds four major imperatives of urban politics: mobilizing an electoral majority, promoting economic growth, maintaining the city's credit, and moderating political conflicts (Shefter 1985, 195). Different sets of institutions, personalities, and circumstances dictate how a regime manages the frequent tension between these goals. Satisfying the demands of one's electoral coalition and abating social conflicts may push spending growth faster than revenues, creating a fiscal crisis. The coalition that mobilizes to "reform" the government and restore the fiscal health of the city (thus serving the other two imperatives) is not, however, electorally broad-based and durable, and thus has difficulty maintaining itself
in the presence of a strong party organization such as exists in New York. Machine politicians may return to office in part by adopting some of the reformers’ agenda (Ibid., 220–221). He describes the cycle during the City’s fiscal crisis:

The regime that governed New York between 1969 and 1975 managed to dampen political conflicts, but did so in a way that was not fiscally viable. It increased benefits to municipal employees and racial minorities without making commensurate reductions in public expenditures benefitting other segments of the city’s population or increasing municipal taxes enough to finance its expenditures. (Ibid., 203)

The successor, post–fiscal crisis regime established a relationship with these groups “that contains conflicts without throwing the city’s budget out of balance” (Ibid., 203). Most importantly, however, the new regime reduced the political power of these groups—especially racial minorities—who previously had seen their power expand.

Shefter’s analysis suggests that the patterns of politics and fiscal policy prevailing in New York and other major cities embody an accommodation between those who control credit and capital, on the one side, and the ideal that government should be subject to popular political control, on the other (Ibid., 235). Although there may be “accommodation,” it has not been stable and institutionalized. There is, as he notes, a “category . . . of cities, such as New York and other major cities in the Northeast, in which the local political system possesses a significant measure of autonomy from the hierarchies of civil society.” The problem of stable fiscal management in New York and other cities is a regime-formation problem. This occurs when the city’s business community has not been actively engaged on a regular basis with public officials and other organized groups who control voting blocs and are thus crucial to the political success of any given administration. Although Shefter uses the term, what he describes is not a regime, in the sense that Stone means it, because it is not cooperative and empowering.

**IS AN ALTERNATIVE TO THE DEVELOPMENT REGIME POSSIBLE?**

There is no doubt that regimes may exist. The question is whether it is possible to have a regime that includes members other than downtown business and financial interests and government. Is it
possible to have a progressive regime, as Rosdil (1991) and Swanstrom (1986) argue? Or, do the Atlanta and Denver cases confirm the virtual inevitability of reliance on a "corporate-center" strategy?

The arguments of the progressive urban agenda proponents have some validity. The very exigency of the space they are able to define, and the very particular conditions under which some flexibility may be available (e.g., Rosdil 1991), demonstrates how much localities are subject to the forces of economic competition. As a result, localities are obliged to adopt pro-development strategies to the detriment of pursuing more "populist" redistributive goals. Swanstrom's policy prescription, for instance, is linkage: downtown office developers must contribute to a fund for low-income housing or for development-related expenses incurred by downtown growth. But the sums involved are small, perhaps $5 million to $7 million annually, in cities with budgets of $800 million to $1 billion (Swanstrom 1986: 104; Bureau of the Census 1982). The experience of these few communities does suggest that a pro-growth environment can coexist with a municipal progressive agenda, but seemingly only in highly particular circumstances, and only to a small degree.

The problem with the Swanstrom/Rosdil analysis is that it fails to accurately locate the principal difficulty that cities actually face. The problem is not some global "economic restructuring," but the political economy of American federalism. Swanstrom touches upon this issue in the conclusion of one of his articles, finding "many of the economic constraints on local governments in the United States are the result of political choices, especially the political structure of American federalism, not the inevitable result of a capitalist economic system" (Swanstrom 1986, 105). But he fails to elaborate on this point and, a few years later, fails to focus on it further (1993). Yet these choices are are not political choices that are available to local officials. The important choices have already been made and can only be unmade at the national and state levels.

It is the political economy of American federalism that lies at the core of the fiscal problems of American cities and goes the furthest toward explaining them. The problems that the political arrangements of American federalism raise and that contribute to the urban fiscal crisis are numerous. Separate suburban jurisdictions encouraged by federal policies and the resulting failure to create a regional
tax base mean that central cities provide services to workers who reside (and pay taxes) in another jurisdiction. Unlike other countries, local governments in the United States must rely principally on private investment and lenders for their capital needs, resulting in competition among cities for development and investment in order to muster the resources necessary to maintain spending. Nor does the federal government assume the burden of redistributive policies, further aggravating the competitive nature of relations among localities (including states), which have little incentive to provide a level of redistributive services greater than their neighbors (Oates 1972; Ladd and Doolittle 1982). The rise and fall of federal and state aid to cities has already been mentioned as another element of the problem, with higher levels of government imposing either unfunded mandates or matching requirements for grants on localities.

The states are the intervening variable between the federal government and local governments. Fuchs's analysis (1982) of why Chicago avoided a fiscal crisis describes Chicago's success in maintaining a narrow functional scope. New York, according to Fuch, could not. There was a kind of regime at work in Chicago, between Mayor Daley and his protegés in Springfield. The result was an intergovernmental structure that kept the city free of a large burden of social spending, unlike New York. It is important to note that Daley was able to help put this structure into place; but not every mayor can arrange such a relationship between city and other competing governmental entities. Daley's success was rooted in the power of his political base.

"Economic determinism" should not be used to replace a different variety of structuralist argument. Most cities have successfully managed fiscal strain through various managerial strategies in order to improve productivity without facing default. Nonethelss, as Ladd (1994) has shown, cities continue to face a situation of scarcity and strain. An addition to this analysis should focus on a neglected variable, the origins of fiscal stress in American cities, which helps explain the dynamic that cities face. In the face of our particular federalist structure, competition among localities remains the most viable option. Consequently, even the national level is constrained by its reliance on private lenders and the mobility of capital. This in turn dictates reliance on a governing coalition that is in conflict with the needs and desires of the electoral
coalition. As has been earlier suggested, the needs of the latter are not addressed adequately by state or federal policies.

Regime theory was developed without focusing on fiscal stress as a preeminent problem. The primary concern was instead managing and directing growth in a dynamic social order. Whether a particular set of relationships can be characterized as a regime depends, in part, on overcoming severe environmental stress. This requires an ability to adapt to new circumstances. One could hypothesize that, at least in some cases, cities that have undergone one or more periods of fiscal stress have failed to build a coalition that bridges the tension between the competing imperatives of the American urban political economy. More research is needed to determine the extent to which the alternative hypothesis presented here—that American federalist arrangements have a large role in contributing to the fiscal stress of cities—is valid.

Finally, there exists a normative dimension to the regime theory perspective on governance. For such a coalition to succeed, all participants must view the process as an opportunity to combine their resources in a way that produces benefits rather than merely divides them. This is in contrast to the conventional political science view of conflict as a zero-sum process. Therefore, the theory presented here suggests that the current direction of federal policy toward decentralization of redistributive functions is likely to increase the problems of local governments, rather than alleviate them.

Notes

2 At least in some cases, cities have actually managed to keep revenues at the same or higher real levels, despite the limits imposed on certain revenue sources and decreases in intergovernmental aid, by diversifying their sources, mostly by increasing user fees and charges. While the overall budgetary level may remain at approximately the same real levels, how the changing revenue sources affect the budgetary allocation among different functions remains an open question. See Niblack and Stan’s (1992) study of Los Angeles.
4 See also Elkin’s description of the governing regime in Dallas.
6See also Reed 1984.

References


Although the Intelligence Authorization Act of 1991 enacted several important statutory requirements for the approval and reporting of covert operations, loopholes remain in the legislation that permit the president to delay notification of covert operations to Congress for extended periods of time and to involve nongovernmental entities in the conduct of such actions. The ambiguity created by such loopholes serves to blur the line between the powers of the executive and legislative branches and to increase the possibility of conflict between the president and Congress in the future. Effective oversight of covert operations in the 1990s will therefore rely heavily on the good faith of both members of Congress and executive officials and on the willingness of both parties to work together in an honest and cooperative fashion.

In Covert Intervention: The Limits of Intervention in the Postwar World (1987), Gregory F. Treverton notes that "secret' operations in a democracy are a paradox" (p. 3). The American political system has traditionally been an open one, in which information is shared freely and the separate branches of government are held account-
able for their actions by an elaborate system of checks and balances. Yet covert action calls for secrecy, planning behind the scenes, and at times unilateral decisions that are made without the benefit of appropriate oversight. It is in this contradiction between openness and secrecy that the dangers inherent in covert action lie. As author Loch K. Johnson has noted, "the enduring irony of intelligence is its potential to destroy as well as to guard democracy" (Johnson 1992–93, 69). In the last twenty-five years much effort has been expended to bring democracy to the intelligence community, to hold it accountable for its actions and decisions, and to establish a system of oversight that will allow the U.S. government to pursue activities vital to its national security while observing basic democratic and constitutional principles.

Former Director of Central Intelligence (DCI) Stansfield Turner has declared that in the 1990s there is "no greater need than to sort out the problems of oversight" (Johnson 1992–93, 67). The Intelligence Authorization Act of 1991 (Public Law 102-88) is the most significant attempt to date to deal with these problems. The act embodies a set of reforms that grew out of the destructive consequences of the Iran-Contra affair. For the first time, the term "covert action" is statutorily defined. In an effort to correct earlier legislative ambiguities, the act also expands the requirements for the approval and reporting of presidential "Findings", which are submitted (usually by the DCI) to Congress to authorize covert actions. It requires that findings be in writing, prohibits these documents from retroactively authorizing covert actions, and calls for them to list all nongovernmental agencies and third parties to be involved in a given operation. It also explicitly prohibits covert actions intended to influence U.S. political processes, public opinion, or media. In doing so, it attempts to mitigate fears that have existed since 1974, when it was revealed that the CIA had engaged in large-scale espionage operations domestically, compiling files on thousands of Americans through Operation CHAOS.

The act, however, is as significant for what it neglects as for what it accomplishes. Like the Hughes-Ryan Amendment and other pieces of oversight legislation before it, the Intelligence Authorization Act allows a loophole for the president to withhold the disclosure of a covert action to Congress for longer than the expected period of 48 hours after the initiation of the operation; he need only report to the Congress in a "timely fashion" (Intelligence Authorization Act 1991, 105 Stat. 443). The ambiguity of this term
allows the president to postpone the disclosure of a covert action for an extended period. The potential for future abuse of the oversight system, like that witnessed in the Iran-Contra affair when President Reagan felt he had the constitutional and governmental authority to postpone the disclosure of the arms for hostages deal indefinitely, is therefore very real.

The Act’s definition of covert action also exempts from the findings requirement any preliminary diplomatic contacts with foreign governments by U.S. government officials that seek to determine the willingness and/or feasibility of their conducting covert actions on behalf of the United States. It therefore allows another loophole through which the president could skirt congressional authority, in this case by permitting a third party that is not accountable to Congress (for example, the security forces of another nation) to conduct covert operations at the president’s request.

Both of these loopholes were added to the act after its original draft (S. 2834) was pocket-vetoed by President Bush on November 30, 1990. Bush viewed S. 2834’s stringent 48-hour reporting requirement as an infringement on the authority granted to his office by the Constitution. He noted that there might be cases in which the extreme sensitivity of an operation would require his office to postpone the submittal of a finding to Congress for more than two days, but that in “almost all instances, prior notice will be possible” (Conference Report 1989, 27). Bush was also concerned with the Act’s statutory definition of covert action, which he felt might “seriously impair” (Conner 1993, 53) the effective conduct of U.S. foreign relations. Bush felt foreign governments would be hesitant to engage in secret negotiations or to conduct covert operations in tandem with the U.S. if they believed that they, too, would be held accountable to congressional oversight procedures. Regardless of the substance or weight of these claims, the Intelligence Authorization Act for Fiscal Year 1991 emerged as a flawed jewel. It enacted several important statutory reporting requirements while preserving certain loopholes and ambiguities that have caused tension and disagreements between the president and Congress in the past. Whether or not the act will permanently eliminate the possibility of “rogue” intelligence community activities and provide for a thorough and cooperative system of intelligence oversight remains to be seen.
In November 1986, press disclosures exposing the secret sale of U.S. arms to Iran called into question the inclination of the executive branch to honor existing oversight arrangements. Further reports that profits from the arms sales had been channeled to the Contras in Nicaragua through Swiss bank accounts in clear violation of an act of Congress (the Boland Amendment) transformed the situation into a full-blown scandal. Members of the House and Senate were outraged that they had to learn about these operations through reports emerging from the Middle East media. The “Era of Distrust” in relations between Congress and the president had formally begun.¹

These initial revelations led to a number of investigations. President Reagan appointed a special review board, headed by former Senator John Tower (R-Texas), to investigate the allegations. Special committees created by the House and Senate held joint hearings on the affair in 1987. It was soon discovered that the covert activities of the Iran-Contra affair were conducted and directed by the staff of the National Security Council (NSC). By removing the operations from the hands of the CIA, the NSC officials involved had excused themselves from observing and abiding by established governmental procedures of oversight and cooperation. It was also revealed that President Reagan knew of the arms sales to Iran (though he denied knowledge of the diversion of arms sales profits to the Contras) and had only retroactively authorized the operation when the CIA insisted on a finding (Crabb and Holt 1989, 186). The finding expressly ordered the DCI not to report the operation to Congress. When questioned about his failure to present Congress with a finding in a timely fashion, President Reagan referred to a 1986 memorandum from the Justice Department’s Office of Legal Counsel, which offered a more liberal interpretation of presidential discretion. Known as the Cooper Memorandum, this document concluded that, because of powers granted to his office by the Constitution and by the 1980 amendments to the National Security Act, a president has “virtually unfettered discretion” (Congressional Record 65 1991, H 2624) in deciding what constitutes timely notice. Many members of Congress reacted sharply to this interpretation, which they felt violated the good faith implicit in existing oversight obligations. As one member later noted, “[The] memorandum was used by the Reagan
administration in an effort to justify its failure to ever provide Congress with the required notice of the covert actions in the Iran-Contra Affair" (Congressional Record 119 1991, H 6160).

The congressional committees investigating the affair noted that it “was characterized by pervasive dishonesty and inordinate secrecy” (House Select Committee to Investigate Covert Arms Transactions with Iran and Senate Select Committee on Secret Military Assistance to Iran and the Nicaraguan Opposition 1987, 13). Vice Admiral John M. Poindexter, National Security Adviser (NSA) during most of the affair, readily admitted to such deceptive practices and noted that “our objective here all along was to withhold information” (Crabb and Holt 1989, 187). The majority report of the congressional committees offered the following conclusion regarding the administration’s abuse of proper procedure:

The findings process was circumvented. Covert actions were undertaken outside the specific authorizations of presidential findings. At other times, covert actions were undertaken without a presidential finding altogether. Actions were undertaken through entities other than the CIA, including foreign governments and private parties. There were claims that the findings could be used to override provisions of the law. The statutory option for prior notice to eight key congressional leaders was disregarded throughout, along with the legal requirement to notify the intelligence committees in a ‘timely fashion.’ (House and Senate Select Committees 1987, 378–79)

As Frank J. Smist Jr. has noted, “such contempt for the American constitutional form of government was unprecedented” (Smist 1990, 264).

In the wake of these revelations, the Reagan Administration moved quickly to upgrade and refine existing oversight procedures. The principal recommendations of the Tower Commission were implemented through a series of National Security Decision Directives (NSDDs). NSDD 266, signed by Reagan on March 31, 1987, stated the intent of his office to assure that “all requirements of law concerning covert activities, including those relating to Presidential authorization and congressional notification, be addressed in a timely manner and complied with fully” (Conner 1993, 47).

NSDD 276, signed on June 9, 1987, established and confirmed NSC committees charged with supervising covert actions in accordance with NSDD 266. NSDD 286, signed by Reagan on October 15, 1987,
banned the NSC staff from conducting covert actions altogether. It also required presidential findings to be in writing and prohibited them from being issued retroactively. The reasons for any delay in the notification of Congress were required to be in writing and would be evaluated by the NSC’s National Security Planning Group every ten days. President Reagan was satisfied with these reforms, stating that they were evidence of his “determination to return to proper procedures, including consultation with Congress” (Conner 1993, 48).

Concerns were still voiced, however, regarding the “timely fashion” language incorporated into existing oversight directives and legislation. Congress soon moved to codify many of the reforms introduced in President Reagan’s NSDDs and to introduce tougher oversight provisions. Two new bills were introduced: S. 1721 in the Senate, and H.R. 3822 in the House. This new legislation signified the intent of many members of Congress and both intelligence committees to limit presidential notification to a strict 48-hour period. Congressional fears stemmed from a belief that presidential notification could be withheld indefinitely as long as the NSC planning group, the NSPG, decided to continue its ten-day evaluations required in NSDD 286. It was clear, however, that any bill that tried to impose a strict time limit on notification would meet with a presidential veto and that the House would not be able to override such a veto. Congress therefore decided to postpone consideration of intelligence oversight legislation until the Bush administration took office. Meanwhile, the use of covert action as a foreign-policy tool was becoming far less commonplace. White House officials noted that the number of operations conducted had been cut in half from 1987 to 1989 (Twentieth Century Fund 1992, 40).

On July 10, 1990, SSCI Chairman Boren introduced the 1991 Intelligence Authorization Bill, S. 2834, on the floor of the Senate. Representative Anthony Beilenson (D-California) introduced the House version of the bill, H.R. 5422, on August 1, 1990. The bills contained measures that would enforce a 48-hour reporting requirement for Presidential findings, statutorily defined covert action, and required the president to notify Congress whenever he requested a nongovernment entity or foreign government to engage in a covert operation on behalf of the United States. Findings were also to be in writing and could not retroactively authorize a covert action.
President Bush, concerned about the new reform proposals of the intelligence oversight committees, forwarded a letter to their chairs on October 30, 1989, in an attempt to clarify his position on the “timely fashion” issue. In the letter he stated that he fully intended to “provide notice in a fashion sensitive to congressional concerns” (Conference Report 1989, 27). He then added, “In those rare instances where prior notice is not provided, I anticipate that notice will be provided within a few days. Any withholding beyond this period would be based upon my assertion of the authorities granted this office by the Constitution” (Conference Report 1989, 27). With these assurances from President Bush, Congress decided to drop the stringent 48-hour reporting requirement from their oversight legislation. Problems still remained, however. The Bush administration did not wish requests to foreign governments to undertake covert actions on behalf of the United States to be treated formally as actual U.S. covert operations, and was therefore skeptical about Congress’ attempts to statutorily define covert action.

Congress, believing that the President was in favor of the bill as a whole despite certain reservations, sent him S. 2834 for signature. Bush refused to sign the bill, however, and it was pocket vetoed on November 30, 1990. In a memorandum of disapproval, Bush stated that he felt the statutory definition of covert action would go too far in restricting U.S. dealings with foreign governments, adding that “the mere existence of this provision could deter foreign governments from discussing certain topics with the United States at all” (Conner 1993, 53).

Congress reacted with surprise at these objections. In a joint explanatory statement that was sent to Bush with S. 2834, it had been stated clearly that the provision was designed “to prevent the conduct of a covert action at the specific request of the United States that bypasses the requirement for administration review, presidential approval, and consultation with the intelligence committees” (Ibid., 54). As one member of Congress noted, “the purpose of the provision . . . was simply to make it clear that the President could not use surrogates to accomplish what he was legally prohibited from doing directly” (Congressional Record 65 1991, H2615). Nonetheless, the intelligence committees soon moved to revise the bill and remove those provisions objectionable to the president.

Meanwhile, as of November 1990, no funding had as yet been approved for the intelligence community for fiscal year 1991. Two new bills were soon introduced, H.R. 1455 in the House in March
of 1991 and S. 1325 in June in the Senate, which contained most of the changes in language and content that the President had requested. Negotiations, however, dragged on for months regarding the specific wording of certain provisions.

Congress found itself losing political ground, however, as the Iran-Contra affair faded from public memory and successes in the Gulf War resulted in strong approval ratings for the Bush administration. Many members of Congress were also in disagreement over the issue of how far a stringent oversight bill would go in affecting the ability of the United States to successfully carry out intelligence activities during periods of conflict. As one legislator noted:

It is my understanding, confirmed in several ways, that had [S. 2834] been enacted into law and not vetoed, Desert Shield and Desert Storm would have been totally different from, and flawed to the extent that perhaps those excellently concluded missions would not have ended in victory for the coalition . . . The wording of that original bill . . . would have hampered American intelligence efforts on a dozen different fronts, including in the diplomatic and military portions of Desert Shield and Desert Storm (Congressional Record 65 1991, H2616).

With Congress divided, the intelligence committees soon acquiesced to the president’s position by placing the “timely fashion” language back into H.R. 1455. In rationalizing their legislative setback, the committees even went as far as to concede that “if the Constitution in fact provides the President authority to withhold notice of covert actions for longer periods, then the conferees’ interpretation cannot be legally binding upon the President” (Conner 1993, 56). On August 2, 1991, H.R. 1455 was presented to President Bush for signature. On August 14, after almost two years of debate and negotiation, Bush finally signed the Intelligence Authorization Act, Fiscal Year 1991. In his statement upon signing the act, Bush reaffirmed his intention to notify Congress within a few days when prior notice could not be given and expressed his confidence that the bill would put to rest disputes that in the past had arisen between Congress and the executive branch. He also stated his belief, however, that certain provisions in the act were unconstitutional and that its statutory definition of covert action was “unnecessary” (Signing the Intelligence Authorization Act 1991, 1138). As he further noted:

Several provisions in the act requiring the disclosure of certain information to the Congress raise constitutional concerns. These
provisions cannot be construed to detract from the President's con-
stitutional authority to withhold information the disclosure of which
could significantly impair foreign relations, the national security, the
deliberative processes of the Executive, or the performance of the
Executive's constitutional duties. (Ibid., 1138)

A mere four years after the Iran-Contra affair, it was thus pains-
takingly obvious that many of the ethical and constitutional issues
the scandal raised had yet to be settled. The legislative loophole
present in the words "timely fashion," which led to so much abuse
of the oversight system during the Reagan administration, had yet
to be closed. Congress and the president also continued to view
their powers and prerogatives within the intelligence community
through very different constitutional and moral lenses.

THE INTELLIGENCE AUTHORIZATION ACT OF 1991:
SCOPE AND EFFECT

The Intelligence Authorization Act of 1991 stands to this day as the
most important piece of legislation affecting congressional over-
sight of covert action. Its reforms represent an attempt to ensure
more executive accountability to Congress by placing oversight
procedures and regulations within a more democratic framework
that stresses the integrity and necessity of our constitutional system
of checks and balances.

Not all of the act's provisions pertain to the oversight of intelli-
gence activities. Title III of the bill contains reforms aimed at
bringing the CIA's retirement systems in line with systems in effect
at other federal agencies. Title IV consists of general provisions,
one of which declares that authorizations in the act are not to be
interpreted as providing authority for intelligence activities not
otherwise authorized "by the Constitution or laws of the United
States" (Ibid., 105 Stat. 434). Title V of the act contains several
provisions of importance to the Department of Defense (DOD).

The most important provisions in the act, however, are con-
tained in Title VI, entitled "Oversight of Intelligence Activities," in
which several measures serve to reform existing intelligence over-
sight law and codify informal reporting procedures. The require-
ments for the reporting and approval of findings are substantially
modified. The act requires that findings be in writing, calls for them
to specify all government agencies and third parties involved in the
conduct of a covert action, bans them from retroactively authoriz-
ing covert actions, and prohibits any findings that authorize actions
that “would violate the Constitution or any statute of the United States” (Ibid., 105 Stat. 443). In addition, it is stated that the president may not authorize a covert action unless he determines that the operation is “necessary to support identifiable foreign policy objectives of the United States and is important to the national security of the United States” (Ibid., 105 Stat. 442).

Congress’ “power of the purse” is affirmed in Section 603 (c) of the act, which states that no funds may be appropriated for, or expended by, any U.S. Government agency or entity for a covert action unless a Presidential Finding has been signed or otherwise issued in accordance with the above requirements. The act also explicitly states that “no covert action may be conducted which is intended to influence United States political processes, public opinion, policies, or media” (Ibid., 105 Stat. 444). Operations similar to those conducted in Operation CHAOS are thereby officially outlawed.

The act greatly clarifies oversight procedures by introducing more concrete and formalized procedures for the reporting and approval of covert actions. The intelligence committees must now be informed of “any significant change in a previously approved covert action, or any significant undertaking pursuant to a previously approved finding” (Ibid., 105 Stat. 443). The DCI is also required to provide periodic statements to Congress explaining the continuing need for ongoing covert operations. Section 501 (e) of the act encourages openness between the legislative and executive branches by stating that “nothing in this Act shall be construed as authority to withhold information from the intelligence committees on the grounds that providing the information to the intelligence committees would constitute the unauthorized disclosure of classified information or information relating to intelligence sources and methods” (105 Stat. 441).

Once again, the president is given the authority to report a finding to a select group of eight members of Congress, consisting of the chairs and ranking minority members of the intelligence committees, the Speaker and minority leader of the House of Representatives, and the majority and minority leaders of the Senate, if he feels that it is “essential to limit access to the finding to meet extraordinary circumstances affecting vital interests of the United States” (105 Stat. 443). The president is also again provided with a loophole through which to postpone notification of a covert
The Intelligence Authorization Act of 1991

action indefinitely. Section 503 (c) (3) of the act states that whenever a finding is not reported pursuant to the requirements listed in the act, “the President shall fully inform the intelligence committees in a timely fashion and shall provide a statement of the reasons for not giving prior notice” (105 Stat. 443). As in prior legislation, what period of time constitutes a “timely fashion” is not considered or defined.

The act also statutorily defines covert action for the first time as “an activity or activities of the United States Government to influence political, economic, or military conditions abroad, where it is intended that the role of the U.S. Government will not be acknowledged publicly” (105 Stat. 443). Certain activities are then exempted from this definition, including traditional counterintelligence activities, diplomatic or military activities, law enforcement activities, and activities to provide routine support to overt activities.

**Effective Oversight or Political Compromise?**

There is little doubt that the Intelligence Authorization Act was an important piece of legislation with important consequences for the future of intelligence oversight. Yet, in view of the concessions granted to President Bush after his pocket veto of the act’s original draft, it is tempting to wonder whether Congress strayed too far from its original position in enacting a compromise on the bill to the detriment of democratic and constitutional principles.

In the fall of 1990, the Twentieth Century Fund convened a distinguished group of scholars and former government officials to consider the future of covert action. The “Task Force on Covert Action and American Democracy” discussed the implications of the Intelligence Authorization Act at length and made several recommendations in its final report as to how oversight reforms in the act can be further upgraded and enhanced in future legislation and presidential directives.

The group concluded that, although covert action is inherently in conflict with the democratic aspirations of our society, “covert action may be justified when a prospective threat creates a compelling national interest that cannot be met prudently by overt means alone” (Twentieth Century Fund 1992, 5). The use of such operations thereby necessitates a set of laws or regulations that serve to minimize the historic tension between covert undertakings and democratic accountability. The task force calls for future reform
efforts in the field of intelligence oversight to establish procedures that ensure effective executive and congressional oversight, "plugging all the gaps revealed five years ago by the Iran-Contra Affair." It is also suggested that a broad consensus be built on the principles under which covert operations may be undertaken. New objectives and methods for covert action will have to be established that address the changing threats and challenges facing the United States in the post–Cold War era.

The task force was critical of the loophole in the Intelligence Authorization Act that allows the president to provide notice in timely fashion in extraordinary circumstances, noting that this "leaves intact a presidential claim of authority to keep Congress uninformed—the very claim invoked in Iran-Contra" (Ibid., 6). It also expresses concern over the Act’s apparent acknowledgement of the legitimacy of using private nongovernmental entities to conduct covert actions. Such entities are frequently accountable to no one, as was shown during the Iran-Contra affair.

The task force's report listed a series of specific recommendations aimed at improving current oversight procedures. In regards to the issue of notification, the task force suggested that, even in cases of extraordinary sensitivity, Congress should be informed of a covert operation within 48 hours of its initiation. In all other instances notification should occur well in advance of the start of an operation, so that the president would receive feedback from Congress on the action's potential pitfalls or weaknesses. The task force also felt that civil and criminal penalties should be enacted for the violation of oversight laws and regulations, including the giving of false or misleading information to Congress. It also strongly recommended that the attorney general, DCI, and secretaries of state and defense provide the president with preliminary statements of concurrence or objection before a Finding is signed. Covert actions should also be used to support only publicly articulated policies and interests.

The task force concluded that while the Intelligence Authorization Act "closes several loopholes that allowed Iran-Contra to happen . . . several significant gaps remain" (Ibid., 10). It notes that the Iran-Contra affair was made possible by private entities with no formal government status or accountability. Under present law, such entities can still play a role in covert activity. As Loch K. Johnson has noted, such privatization of covert actions invariably
The Intelligence Authorization Act of 1991 has the effect of removing intelligence policy from a system of proper constitutional checks and balances (Johnson 1989, 254).

Some critics of modern congressional oversight have focused more closely on the constitutional issues involved in the division of power between the legislative and executive branches in the realm of foreign policy. Many supporters of presidential supremacy in the foreign-policy arena refer to Article II, Section 1, in our Constitution which declares that the president shall have responsibility for our nation’s foreign affairs. They also note the writings of Thomas Jefferson, who believed that the transaction of business with foreign governments should be executive altogether.

Proponents of congressional checks on executive power disagree with this interpretation. Michael J. Glennon believes that the constitutionality of congressional oversight of intelligence activities rests on the same framework applicable to other executive activities, such as the war power and executive agreements. Glennon notes that the case for the constitutionality of restraints on covert operations is further strengthened by the fact that “Congress, not the President, created the CIA and assigned its functions and responsibilities” (Glennon 1990, 304). Congress should therefore hold the power to decide when the president can or cannot use the facilities and resources of the CIA to conduct covert actions. As Glennon tells us, “even the minority report of the Iran-Contra committees acknowledged the constitutionality of a statute that would ‘set rules for the use’ of funds Congress makes available to the intelligence community” (Ibid., 304). He then refers to the Supreme Court’s decision in Youngstown Sheet & Tube Co. v. Sawyer, which reinforced Congress’ prerogative to oversee executive activities in the field of foreign relations. Glennon points to the remarks of Professor William Van Alstyne who, in looking at the decision, concluded that

to the extent Congress desires clarity, certainty, and reliability in highly problematic areas of executive direction of foreign intelligence activity, it is both appropriate and essential that Congress should say so, by law, as concretely as it can agree to do. (Ibid., 306)

Glennon’s own view is that “if we are truly a government of laws, then laws limiting [the CIA’s] discretion, and not merely words of good will uttered by executive officials at closed congressional hearings, are absolutely essential” (Ibid., 303). Such laws must place the “most rigorous limitations” (Ibid., 304) on covert operations.
Most observers would agree, however, that any oversight law or procedure is largely impotent if it is not observed in good faith by both the executive and legislative branches. As Frank J. Smist has observed, “truth is the vital ingredient that enables oversight to work” (Smist 1990, 265). In Smist’s mind, such honesty and forthrightness was the main casualty of the Iran-Contra affair. Vice Admiral Poindexter personified the unwillingness of executive officials to cooperate with congressional oversight when he declared that he withheld information because he “simply did not want any outside interference” (Johnson 1989, 128).

Trust and good faith are two-sided coins, however, and executive officials are far more willing to cooperate with intelligence committee members when they feel that congressional members can be trusted with sensitive national secrets. Executive distrust of Congress has often led to calls for the creation of a joint congressional intelligence committee, where fewer individuals would have to be briefed by CIA officials. As Senator John W. Warner (R-Virginia) has noted, such proposals are “rooted in the belief that Congress cannot keep information secret” (Warner 1989, 108). While Warner refutes this assumption by declaring that most information leaks actually come from the executive branch, he recognizes that its mere existence has been a historic source of tension between the president and Congress (Ibid., 108). It is therefore apparent that the effectiveness of the Intelligence Authorization Act and any future oversight legislation will be determined largely by the level of cooperation between the legislators and executive officials charged with implementing it.

**CONCLUSION**

During the congressional debate on the Intelligence Authorization Act, one member of Congress noted, “we need laws on the books that erase ambiguities and set up clearly defined procedures” (Congressional Record 65 1991, H 2614). While the Intelligence Authorization Act does enact several important statutory requirements for the approval and reporting of covert operations, a “zone of twilight” (Warner 1989, 109) still seems to exist between where the president’s powers cease and those of Congress begin. Lingering ambiguities seem to center around the loopholes in the act that allow the president to delay notification of covert operations to Congress for extended periods of time in certain cases and to
The Intelligence Authorization Act of 1991

involve private nongovernment entities in the conduct of such actions.

Specific time-limit legislation was incorporated into S. 2834 and would have become law if not for President Bush’s pocket veto in late 1990. After the onset of Operation Desert Storm, Congress seemed to lose much of the political capital it had acquired after the Iran-Contra affair. Time-limit guidelines were therefore never enacted in the Intelligence Authorization Act, which leaves intact the language of “timely fashion.” The pressing need for specific time-limit legislation had been described eloquently by Senator John W. Warner:

The Constitution gives the power to appropriate funds to Congress and, by statute, that appropriated funds shall be expended only for purposes authorized by Congress. Because Congress authorizes and appropriates funds, it has, according to the Courts, the implied power to oversee all government spending and seek such information as necessary to make informed decisions. In order to exercise this authority over covert actions, the intelligence committees must be informed at the earliest possible moment (Warner 1989, 108).

If such legislation proves difficult to enact, there are fallback positions preferable to those taken by Congress after the pocket veto of S. 2834. A compromise might be reached, as Lloyd Cutler has suggested, in which the president informs the intelligence committees straightaway: “I have signed a Finding” and because of extraordinary circumstances a full report on the activities involved will have to be delayed (Johnson 1989, 253–4). This option at least allows Congress to track delays in presidential notification and avoid embarassing situations in which it has to be informed of covert operations by executive leaks or the media.

Equally disturbing is the fact that private nongovernmental entities can still be used to conduct covert activities, providing that the president provides Congress with a list of all the third parties he forsees engaging in a given operation. Actors that are not directly accountable to any branch of government will therefore continue to be trusted with our nation’s most sensitive secrets and with the responsibility for carrying out operations vital to the national security of the United States. The reality of this situation became a nightmare in November of 1986, when media reports concerning the Iran-Contra affair began. That the possibility of another such scandal occurring exists seven years later is shocking.
In the twenty years since the passing of the Hughes-Ryan Amendment, intelligence oversight has come a long way. The Intelligence Authorization Act, Fiscal Year 1991, is a symbol of Congress' continuing determination to address its oversight responsibilities seriously. The act has gone far to clarify the president's responsibilities in reporting covert operations to the legislative branch. Many ambiguities were also cleared up in the act, such as the precise definition of covert action. But the act retains many loopholes that blur the lines between congressional and executive authority and allow for the possibility of future conflict. The willingness of executive officials to work in a cooperative and honest fashion with their legislative counterparts will therefore play an important role in determining how capable the Intelligence Authorization Act is of providing effective oversight. As Loch K. Johnson has noted, it is important for executive officials to abide by the policy of "the four cs" in their briefings of Congress: consistency, candor, completeness, and correctness (Johnson 1992–93, 67–68). Only in this manner can Congress and the executive once again enter an "Era of Trust," this time on an equal constitutional footing and with cooperation as a respected goal.

Notes
2. Among the more notable members of the group were Richard E. Neustadt, Gregory F. Treverton, Admiral Stansfield Turner, Theodore C. Sorenson, Lloyd N. Cutler, and John C. Whitehead.

References
This paper examines the effectiveness of linking China's most-favored-nation (MFN) status to economic, foreign-policy, and human-rights issues. It calculates the Sino-U.S. trade balance and reviews two studies on the economic impact of suspending China's MFN status, arguing that loss of MFN would have a major impact on the Chinese economy and a smaller but significant effect on the U.S. economy. The paper examines efforts to link China's MFN status to other issues and concludes that MFN has been most effective in securing economic concessions, less effective in achieving foreign-policy objectives, and only marginally effective in securing human rights. The paper argues that the absolute costs to the United States of MFN suspension politically outweigh the larger relative costs to China, limiting the credibility of threats to suspend China's MFN status. Instead of relying on MFN as a policy instrument, the United States should develop a more selective set of sanctions to pressure China on economic, foreign-policy, and human-rights issues.

This paper examines the effectiveness of linking the renewal of China's most-favored-nation (MFN) status to economic, foreign-policy, and human-rights issues. After analyzing economic aspects of MFN, the paper lists the variety of economic, foreign-policy, and

human-rights issues that have been explicitly or implicitly linked to MFN renewal and assesses the relative success of linkage in different issue areas. The paper concludes that MFN has been most effective in securing economic concessions, less effective in achieving foreign-policy objectives, and only marginally effective in securing human rights. The paper further argues that MFN is too blunt an instrument to form the basis of U.S. policy toward China and that the threat of nonrenewal is becoming less credible and imposing increasing costs on bilateral ties. The U.S. should develop a more selective set of sanctions to pressure China on foreign-policy and human-rights issues.

THE LEGAL BASIS OF MOST-FAVORED-NATION STATUS

Most-favored-nation status refers to the tariff rates charged on imports. In the United States, countries with MFN status are subject to Column One tariff rates, which reflect the seven rounds of General Agreement of Trades and Tariffs (GATT) tariff reductions prior to the latest Uruguay Round reductions. Afghanistan, Cambodia, Cuba, Laos, Montenegro, North Korea, Serbia, and Vietnam are the only countries currently denied MFN status (IBERC 1993, 5). Imports from countries without MFN are subject to Column Two tariff rates, which generally date back to the Smoot-Hawley Tariff Act. The average tariff rate under Smoot-Hawley in 1930 was 44.9 percent; the average tariff rate on dutiable U.S. imports in 1992 was only 5.2 percent (IBERC 1993, 7).

U.S. tariff rates typically include a per-unit duty and a percentage of the total value of the imported goods. Effective tariff rates therefore depend on both the volume and value of trade in specific Harmonized Tariff Categories; if the composition of trade shifts, the overall effective tariff rate will also change. Based on the composition of 1992 imports from China, the average duty rate with MFN was 7.9 percent. Without MFN, tariffs would have averaged 46.1 percent (IBERC 1993, 8). Based on the composition of 1990 imports, the average MFN rate was 8.8 percent and the non-MFN rate would have been 50.5 percent (Solomon 1990). Loss of China's MFN status would significantly increase the cost of Chinese products to U.S. consumers and greatly reduce the volume of Chinese imports. (The economic impact is considered in more detail later.) Table I shows effective MFN and Column 2 tariff rates for the top ten U.S. imports from China (IBERC 1993, C1–C3).
Table I: MFN and Column 2 Tariff Rates, Top Ten U.S. Imports from China, 1992

<table>
<thead>
<tr>
<th>HTS Item</th>
<th>Description</th>
<th>MFN duty</th>
<th>Column 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>64029915</td>
<td>Certain plastic footwear, not covering the ankle</td>
<td>6.0%</td>
<td>35.0%</td>
</tr>
<tr>
<td>61109000</td>
<td>Silk Blend Sweaters</td>
<td>6.0%</td>
<td>60.0%</td>
</tr>
<tr>
<td>95039060</td>
<td>Miscellaneous toys</td>
<td>6.8%</td>
<td>70.0%</td>
</tr>
<tr>
<td>64039990</td>
<td>Certain female leather footwear</td>
<td>10.0%</td>
<td>20.0%</td>
</tr>
<tr>
<td>95034110</td>
<td>Stuffed toys</td>
<td>6.8%</td>
<td>70.0%</td>
</tr>
<tr>
<td>27090020</td>
<td>Crude Oil</td>
<td>$.105/barrel</td>
<td>$.210/barrel</td>
</tr>
<tr>
<td>62061000</td>
<td>Female woven silk blouses</td>
<td>7.5%</td>
<td>65.0%</td>
</tr>
<tr>
<td>64039960</td>
<td>Certain male leather footwear</td>
<td>8.5%</td>
<td>20.0%</td>
</tr>
<tr>
<td>67029035</td>
<td>Artificial flowers</td>
<td>9.0%</td>
<td>71.5%</td>
</tr>
<tr>
<td>03061300</td>
<td>Frozen shrimp</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Article I of the General Agreement of Trades and Tariffs specifies that MFN tariff rates will be applied to all GATT members; extension to nonmembers like China depends on the discretion of the GATT member. The United States used its discretionary power to provisionally grant MFN status to China in 1973. China and the United States formally granted each other MFN status in the U.S.-China trade act signed in mid-1980. Since China was a nonmarket economy, MFN status was subject to the provisions of the Jackson-Vanik amendment to the Trade Act of 1974. The Jackson-Vanik amendment, originally intended to encourage the Soviet Union to allow free emigration of Jews, required the president to certify annually that the government of a nonmarket economy either granted freedom of emigration to its citizens or that extension of MFN status would encourage free emigration (Harding 1992, 96). Annual certification provided Congress a means of evaluating and influencing the president’s policy toward Russia; the recertification process also became the key forum for debate over the president’s China policy after 1990. Emigration was not initially an important issue in Sino-U.S. relations; Deng Xiaoping offered to send the United States as many immigrants as it wanted (Harding 1992, 96). China’s most-favored-nation status was routinely renewed until Chinese troops massacred student protesters in Tiananmen Square on June 4, 1989.
U.S. RESPONSES TO TIANANMEN SQUARE

The Bush administration responded to Tiananmen Square by suspending military sales to China and canceling high-level military and diplomatic exchanges. Bush also allowed Chinese students to remain in the United States, postponed loans to China from international financial institutions, and suspended a range of technology transfers. Bush, who served as the American envoy to China in the mid-1970s, was conscious of China's strategic value and determined to preserve the Sino-U.S. relationship. A few members of Congress called for immediate revocation of China's MFN status, but the bill Congress passed in July 1989 only codified and slightly extended the sanctions Bush had already imposed. Members of Congress gradually became frustrated by the negligible impact of sanctions on Beijing's human-rights policies and by the Bush administration's reluctance to press harder for political reforms. In 1990 the House passed a bill revoking China's MFN status as well as a competing measure conditioning MFN renewal on China's progress toward specific political reforms, but neither bill won Senate approval (Harding 1992, 247–269). In both 1991 and 1992 Congress succeeded in passing legislation establishing conditions for the renewal of China's MFN status. Bush vetoed the legislation each time, preferring to rely on measures like technology restrictions and Super 301 cases which had a more limited impact on Sino-U.S. relations.

U.S. policy toward China became an issue in the 1992 presidential election, with Clinton accusing Bush of coddling China and promising to impose tougher human-rights conditions on the renewal of MFN. Once in office, President Clinton signed an Executive Order on May 28, 1993, which renewed China's MFN status unconditionally for 1993 but also imposed a detailed set of human-rights conditions for renewal in 1994. On May 26, 1994, President Clinton announced that even though Chinese human-rights performance fell short of the Executive Order's standard of "overall, significant progress," he had decided to renew China's MFN status and to delink human rights from future determinations of China's MFN status (Clinton 1994). With this framework in mind, this paper will now consider the potential economic impact of withdrawal of China's MFN status.
Several methodological issues arise when analyzing Sino-U.S. trade. One of the most important issues is whether to use Chinese or U.S. customs statistics. Each country's official figures show that their country is running a deficit in bilateral trade! (Tables II and III).

**Table II: U.S.-China Trade (PRC Customs Statistics, US$billions)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese Exports</td>
<td>3.00</td>
<td>3.40</td>
<td>4.40</td>
<td>5.20</td>
<td>6.159</td>
<td>8.594</td>
</tr>
<tr>
<td>Chinese Imports</td>
<td>4.80</td>
<td>6.60</td>
<td>7.90</td>
<td>6.60</td>
<td>8.008</td>
<td>8.900</td>
</tr>
<tr>
<td>Chinese Trade Deficit</td>
<td>-1.80</td>
<td>-3.20</td>
<td>-3.50</td>
<td>-1.40</td>
<td>-1.849</td>
<td>-.306</td>
</tr>
</tbody>
</table>

**Table III: U.S.-China Trade (U.S. Customs Statistics, US$billions)**

|-------|------|------|------|------|------|------|------|

Both countries calculate exports based on FOB (free on board) and imports based on CIF (cost, insurance, and freight), resulting in a 10 to 15 percent discrepancy in figures. The most important difference lies in the treatment of goods shipped through Hong Kong. The United States counts Chinese goods shipped through Hong Kong as Chinese exports, but counts American goods shipped through Hong Kong to China as exports to Hong Kong. Conversely China counts most American goods shipped through Hong Kong as imports from the United States and counts Chinese goods shipped to the United States through Hong Kong as exports to Hong Kong (Wang 1993, 456). In 1990 the discrepancy between U.S. and Chinese figures for Chinese exports to the United States was $10.037 billion, a figure that closely corresponds with official Hong Kong figures of $10.500 billion of Chinese goods reexported to the United States (The World Bank 1994, 265). Since approximately 70 percent of Chinese exports to the United States are transshipped through Hong Kong, this is the largest source of the discrepancy between trade figures and a reason to prefer U.S. figures (U.S.-China Business Council 1992, 16).
Some analysts (IBERC 1993, 42–43) argue that the rapid growth of direct investment in China from Hong Kong and Taiwan and the transfer of production facilities to the mainland make U.S. trade balances with “Greater China” a better measure of the true trade deficit. Chinese officials estimate that some 36 percent of Hong Kong’s industry has moved across the border to Guangdong province (Baldinger 1992, 14).

**Table IV: U.S.-Greater China Trade (U.S. Customs Statistics, US$billions)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. Exports to Greater China</td>
<td>14.893</td>
<td>22.837</td>
<td>23.381</td>
<td>23.114</td>
<td>27.597</td>
<td>31.745</td>
</tr>
<tr>
<td>U.S. Imports from Greater China</td>
<td>40.769</td>
<td>43.463</td>
<td>46.023</td>
<td>47.525</td>
<td>51.271</td>
<td>60.117</td>
</tr>
</tbody>
</table>

Table IV indicates that the U.S. trade deficit with Greater China has not increased dramatically over the last five years. This is largely because U.S. exports to Hong Kong and Taiwan have increased fast enough to cover most of the dramatic growth in imports from China. (U.S. imports from Hong Kong and Taiwan declined slightly over this period). Aggregating the data hides the fact that Chinese import restrictions have limited U.S. exports to China, a factor relevant for the bilateral trade relationship. Moreover, loss of MFN status would only affect Chinese-origin goods, making trade with China the appropriate measurement. The remainder of this paper will use U.S. Customs figures.

**Costs of MFN Suspension to China**

This paper critically evaluates two studies that calculate the effect that loss of MFN status would have on Chinese exports to the United States. The first is a World Bank study using the World Bank-UNCTAD SMART trade projection model (The World Bank 1994, 155–58, 252–53). The study examined the top 15 four-digit SITC imports (which made up just over half of 1990 U.S. imports from China) and calculated the change in trade volumes caused by imposing Column Two tariffs on Chinese imports. The study
considered two cases: Scenario I allocated the trade destruction caused by higher import prices across all countries, and Scenario II allocated all the trade destruction to China.

The second is a 1993 study by the International Business and Economic Research Corporation (IBERC) using the U.S. International Trade Commission CADIC model (IBERC 1993, B1–B18, C1–C7). The IBERC study considers two primary cases, with three variations on each. The first case examines the impact on the top 25 Harmonized Tariff System (HTS) import categories, which made up 33.9 percent of 1992 U.S. imports from China. This tends to underestimate the impact on trade, since the tariff increases on 2 of the top 25 goods (crude oil and frozen shrimp) are 0.6 percent and 0 percent respectively. The second case considers HTS chapters selected on the basis of potential change in trade volume, which tends to overestimate the impact on trade. The selected chapters make up 71 percent of 1992 U.S. imports from China. The IBERC study also considers the impact of Chinese absorption of the higher tariffs in an effort to maintain export markets. In the 100 percent pass-through case, U.S. consumers pay the full cost of the higher tariffs; in the 75 and 50 percent cases, Chinese exporters bear some of the costs of the higher tariffs in order to maintain exports. Both studies extrapolate from the subset of imports modeled to obtain estimates of the total impact of the loss of MFN.

Table V shows the decline in Chinese imports under each study. The shaded cells represent calculations in the original study; the other cells are extrapolations based on the estimated percentage decline in the volume of trade multiplied by the dollar value of Chinese imports in the given year. These extrapolations introduce two additional sources of error: (1) they ignore changes in the composition of imports from China, and (2) they assume that elasticities are constant as the volume of trade changes (since both studies use isoelastic estimates for the elasticities of demand, supply, and substitution, the second source of error is probably not significant). Nevertheless, the extrapolations allow comparisons between the two studies and estimate the impact of the loss of MFN based on 1993 trade volumes.
Table V: Impact of the Loss of MFN on U.S. Imports from China (US$billions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>World Bank</td>
<td>15.237</td>
<td>18.969</td>
<td>25.728</td>
<td>31.530</td>
</tr>
<tr>
<td>HTS chapters, 75%</td>
<td>-74%</td>
<td>-11.275</td>
<td>-14.037</td>
<td>-19.039</td>
</tr>
<tr>
<td>IBERC Top 25</td>
<td>-34%</td>
<td>-5.120</td>
<td>-6.374</td>
<td>-8.645</td>
</tr>
</tbody>
</table>

These figures give a range of values for the probable decline in U.S. imports from China. The IBERC figures are preferable to the World Bank estimates, because the errors induced by extrapolation to 1993 are lower. The IBERC study suggests the 75 percent pass-through case is the best estimate of the impact of the loss of MFN. Based on 1992 data, the IBERC study suggests Chinese exports to the U.S. would fall by between $12.658 and $19.039 billion. Based on 1993 data, the IBERC study suggests Chinese exports to the U.S. would fall by between $15.513 and $23.332 billion.

A decline in exports to the U.S. is not a total loss; goods can be sold elsewhere and production resources can be reallocated. The partial equilibrium models reviewed above cannot estimate the overall impact of the loss of MFN status on the Chinese economy. A large decline in exports would have ripple effects throughout the economy, significantly slowing China’s growth rate. Sun Shangqing, a Chinese economist, claims loss of MFN would seriously affect only 10 percent of China’s exports, most of which could be
redirected to Asian markets, and that China would be able to meet its 9 percent growth target even with the loss of MFN (Tyler 1994). However a recent World Bank study found that across-the-board loss of MFN could slash China’s economic growth from 13 percent in 1993 to as little as 6.5 percent (Chen 1994). Yizi Chen, an economist who left China after Tiananmen Square, notes “Chinese officials are trying to underplay the impact on its economic development for political reasons.” Chen estimates that withdrawal of MFN status would result in the loss of more than 5 million Chinese jobs, a figure far higher than the Chinese government’s unofficial estimate of 2 million (Chen 1994).

The political impact of this unemployment is unpredictable, but loss of MFN status might not promote political reform. The Chinese government is likely to become more repressive if an economic slowdown causes political unrest. The hardest hit Chinese exports would be toys, shoes, and apparel, sectors in which the increase in tariff rates is high and China commands a large share of the U.S. market. The short-term impact would be greatest on the factories producing for the U.S. export market; these tend to be the private and collective firms in the southeastern coastal provinces that have been driving China’s economic growth. Many of these firms are financed by capital from Hong Kong, Taiwan, or (in some cases) the United States. If MFN was expected to remain suspended, production of many goods would begin to shift back to Hong Kong or Taiwan.

**Costs of MFN Suspension to the United States**

Removing China’s MFN status would impose substantial costs on U.S. consumers. The IBERC study estimates the 1992 increase in consumer cost (in the 75 percent pass-through case) at $13,395 billion using the top 25 HTS items and $16,001 billion using the selected HTS chapters. These figures are obtained by estimating the higher prices and lower sales volumes using the model, resulting in cost increases of $4.464 billion using the top 25 HTS items and $5.334 billion using the selected HTS chapters. The IBERC study scales these figures up by a factor of three “to incorporate the standard assumption of markup in price from the first or landed cost to the price at the retail or final consumption level” (IBERC 1993, 12). Scaling may be appropriate, but the IBERC study fails to note that these transfers take place entirely within the United States,
representing a transfer from U.S. consumers to U.S. producers and distributors. Consumer costs would rise most on items such as toys, shoes, and artificial flowers, with increases of as much as 22.4 percent (IBERC 1993, B15). One lobbyist described the higher toy prices as “a lump of coal in the stockings of kids all across America” (Sands 1994).

The increase in consumer costs must be balanced against increased tariff revenues and additional producer surplus for domestic producers and distributors. The additional tariff revenues obtained by removing China’s MFN status could be estimated by applying the average increase in duty to the remaining trade volume. This would overstate the increase in revenues, because trade will drop most for those goods that have the highest increase in tariffs. This paper calculates tariff revenues by multiplying the remaining imports from China in specific sectors by the higher tariff rates in that sector. Old tariff revenue is then subtracted from new revenue to get the change. Using the top 25 HTS items and 75 percent pass-through assumptions, tariff revenue from China increases by an estimated $2.425 billion. Using the selected HTS chapters, tariff revenue increases by only $0.125 billion, but this underestimates revenue because these chapters were selected on the basis of a large change in trade volume.

Some additional tariff revenue is also generated by increased imports from third countries, but since this revenue is taxed at lower MFN rates and the IBERC study does not include these values, this revenue has been omitted. These lost revenues are more significant for the selected HTS chapters, since more trade is replaced by foreign imports in this case. Higher prices for domestic producers represent an increase in domestic producer surplus, which cannot easily be calculated from the IBERC data. Additional deadweight loss from higher tariffs has also been omitted. The unscaled increase in consumer cost and changes in tariff revenues allow an approximation of the direct costs of MFN suspension to the United States (Table VI). Due to the inability to calculate precisely additional tariff revenues from foreign countries and additional producer surplus, these estimates are likely to be high; the estimate based on the top 25 HTS items is likely to be more accurate.

These estimates ignore the impact of Chinese retaliation against U.S. exports. U.S. exports are granted most-favored-nation status
Table VI: Estimated Direct Costs of MFN suspension to the United States (US$billions)

<table>
<thead>
<tr>
<th>Top 25 HTS items</th>
<th>Selected HTS chapters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher consumer costs</td>
<td>4.464</td>
</tr>
<tr>
<td>Tariff revenues</td>
<td>-2.425</td>
</tr>
<tr>
<td>Producer surplus</td>
<td>unknown</td>
</tr>
<tr>
<td>Dead weight loss</td>
<td>unknown</td>
</tr>
<tr>
<td>Total loss to U.S.</td>
<td>2.039</td>
</tr>
</tbody>
</table>

in China, although the value of MFN is reduced due to high Chinese tariffs and import restrictions. China would likely retaliate by removing American MFN status, but no estimates of this impact are available. Additionally, and more importantly, China could use import restrictions and purchasing decisions to maximize the domestic political costs of denying MFN. Other countries offer ready substitutes for fertilizer, wheat, and most chemical products; China would probably shift purchases of these commodities to other countries to hurt U.S. exporters. Aircraft purchases are the single largest U.S. export to China; Boeing officials expect China to buy $40 billion in new aircraft over the next 15 years (Garan 1994). China would almost certainly publicly target purchases of U.S. aircraft in order to make President Clinton pay a large political price for denying MFN status. China is also a large and growing market for telecommunications equipment. Robert E. Allen, CEO of AT&T, called for unconditional renewal of MFN, citing fierce competition with foreign rivals for part of China’s massive telecommunications market. He notes that “China is becoming the largest market in the world for almost any product you can name” (Harbrecht and Borrus 1994). Major U.S. exports to China are listed in Table VII (IBERC 1993, 16–17).

Some businessmen estimate that China would retaliate for loss of MFN by halting almost all imports of U.S. products (IBERC 1993, 15). However China is likely to continue to import raw materials and machine tools needed for exporting industries while carefully targeting vulnerable sectors of the U.S. economy to maximize the political impact of its response. Its political objective would be to regain MFN status while making a minimum of concessions, a goal that would tend to temper Chinese retaliatory measures. Retaliation might reduce the $8.760 billion in 1993 U.S. exports to China by 50
percent or more. This would have a significant impact on U.S. jobs. Using the Commerce Department’s rule of thumb that each billion dollars of exports supports 19,100 U.S. jobs, a 50 percent reduction in exports could cost 83,658 jobs; complete elimination of U.S. exports to China could cost 167,316 jobs (IBERC 1993, 27). In addition, U.S. businesses have invested nearly $6 billion in China (IBERC 1993, 32). U.S. investments that bring technology transfers to China are likely to be unaffected, but investments devoted to producing consumer goods or other products for the Chinese market could be subject to administrative harassment and denied access to raw materials.

Table VIII summarizes the estimated costs of mutual suspension of MFN status and probable Chinese retaliatory measures.
Export losses and direct costs are not comparable; comparisons must be made within categories. Table VIII indicates that although both countries would suffer from suspending MFN, Chinese costs would be considerably higher than U.S. costs.

An additional consideration is the impact on Hong Kong. Seventy percent of Chinese exports to the United States are transshipped through Hong Kong. The British government estimated in 1994 that the loss of China’s most-favored-nation status could cost the colony $17 to $25 billion in trade and $2 to $3 billion in income, with an estimated loss of 54,000 to 75,000 jobs. Christopher Jackson, the Deputy Director-General of Trade, noted that a reduction of up to $9.936 billion in reexports was double the 1990 estimate, reflecting the rapid growth in Hong Kong’s reexports of Chinese-made goods to the United States (Hong Kong Economic and Trade Office 1994). Hong Kong Governor Chris Patten has stated, “Because of the intimate, symbiotic relationship between Hong Kong and the Chinese economy, anything which damages the Chinese economy, anything which restricts trade between our two most important trading countries, is potentially bearing damage for us.” Patten estimated that loss of China’s MFN status could cut Hong Kong’s economic growth in half, and warned that “trade is a weapon with two edges, you can’t cut the other person without cutting yourself” (Agence France Presse 1993).

### Table VIII: Estimated Costs of Mutual Suspension of MFN, 1993 (US$billions)

<table>
<thead>
<tr>
<th></th>
<th>China</th>
<th>United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost Exports</td>
<td>$15.608-$23.476</td>
<td>$4.335-$8.670</td>
</tr>
<tr>
<td>Direct Costs</td>
<td>Unknown, but high</td>
<td>$2.039-$5.209</td>
</tr>
<tr>
<td>Lost Jobs</td>
<td>2 million jobs</td>
<td>83,658 jobs</td>
</tr>
</tbody>
</table>

**THE CLINTON ADMINISTRATION’S SHIFT IN POLICY**

Recognizing the high domestic and international costs of withdrawing MFN, the Clinton administration gradually backed away from the policy expressed in the May 1993 executive order before ultimately deciding to delink China’s MFN status and human rights. President Clinton reportedly told aides and lawmakers in March 1994 that withdrawing China’s trade benefits was the last thing he wanted to do, but that he felt politically, morally, and legally bound to fulfill the basic terms of the executive order (Friedman 1994b).
Treasury Secretary Lloyd Bentsen and Commerce Secretary Ronald Brown spoke out against the policy of linking China's MFN status to its human-rights record during a March 22 meeting at the White House (Harbrecht and Borrus 1994). A senior official stated that the administration told the Chinese, "Get us over the executive order, and we will consider all kinds of alternatives" (Friedman 1994b). When Secretary of State Warren Christopher visited China in March to lay out the conditions for renewal of MFN, Chinese officials gave him a cool reception and arrested several prominent dissidents to signal their displeasure at U.S. tactics. Some observers interpret Chinese behavior during Christopher's visit as a sign of confidence that the United States could not afford to withdraw MFN status regardless of Chinese compliance with the executive order's conditions. A State Department official admitted that the administration examined selective sanctions in order to "signal that we're looking for a way to pull the trigger" on trade sanctions, an indication that withdrawal of MFN status was no longer considered a credible threat (Lippman 1994).

Clinton administration officials reportedly considered a variety of narrow, more selective sanctions with the goal of pursuing human rights while preserving private enterprise and American investment in China (Greenhouse 1994). These included proposals for sanctions on industries and goods where most Chinese production comes from state-owned enterprises as well as a ban on imports of Chinese-made assault rifles (Harbrecht and Borrus 1994). House majority leader Richard Gephardt suggested that China's MFN status should be renewed with conditions such as revoking trade benefits for state-owned Chinese industries while allowing benefits to continue for private enterprises. Gephardt argued that the goal of a policy of selective sanctions should be to keep the pressure on Beijing to permit greater human rights "but not totally removing MFN so there can be no trade." Representative Nancy Pelosi, who has advocated revoking China's MFN status since 1989, suggested that targeted sanctions were better than no sanctions at all. Pelosi advocated "revocation that is targeted to products made by the Chinese military" and "targeting other products which are for the most part made by the state sectors." She noted that "the U.S. has an opportunity to initiate sanctions which impact the Chinese government the most, the Chinese private
sector to a lesser degree, and the American consumer the least” (Friedman 1994a).

Government officials argued that targeted sanctions could be a legal and administrative nightmare to enforce (Friedman 1994a). Sanctions targeting state-owned firms and military production while allowing similar goods produced by private companies to be imported would invite fraud and “cooperative ventures” between private and state firms to disguise the origin of goods. Some trade experts regarded selective sanctions as totally unenforceable, questioning how American customs officials could ever determine whether a dress or toy made in China was made by a state enterprise or a private enterprise (Greenhouse 1994). Sanctions targeting specific goods or sectors dominated by state and military enterprises to avoid this problem would inevitably hurt private firms to some degree. Targeted sanctions would also be likely to provoke Chinese retaliation, but this would probably be proportional to the size of U.S. sanctions. Ultimately, the Clinton administration decided not to impose selective economic sanctions based on the Chinese government’s human-rights performance, opting only for a ban on imports of Chinese munitions (Clinton 1994).

**Evaluating the Effectiveness of Linkage**

Because MFN was viewed as the best leverage the United States had against China, it was implicitly or explicitly linked with a wide variety of human rights, trade, and international issues. Linkages were conveyed through a variety of official and unofficial means: direct communications through diplomats, official statements of policy, administration comments to the press, and Congressional opinions. A visit by the secretary of state has a different impact than one congressman’s offhand comment that MFN should not be renewed unless the Chinese grant freedom to Tibet. The implied threat of removal of MFN status may have encouraged Chinese concessions on other issues, but it is difficult to be sure when the linkages are so indirect. Moreover, many Chinese concessions were the result of bilateral discussions not directly related to MFN.

With these caveats in mind, this paper will review 1990 and 1991 Chinese concessions that may have been linked to MFN and then assess Chinese performance against the standards set by Clinton’s May 28 executive order. Congressional legislation in 1990, 1991,
and 1992 proposed a wide variety of conditions on the renewal of China's MFN status, including (Harding 1992, 268–69, 278):

- the release of prisoners arrested in the Tiananmen crisis
- greater freedom of the press
- an end to jamming of Voice of America broadcasts
- significant progress toward ending religious persecution
- loosening of travel and study-abroad restrictions
- a ban on the export of goods produced with prison labor
- an end to coercive abortions and sterilizations
- unequivocal commitments to nuclear and ballistic missile nonproliferation
- adherence to the joint declaration on the future of Hong Kong
- a ban on the supply of arms to the Khmer Rouge
- significant progress toward human rights
- correction of unfair trading practices

Although the Senate did not vote on the House resolutions passed in 1990 and President Bush vetoed the bills passed in 1991 and 1992, the items listed above give an indication of the range of issues that Congress felt should be linked to MFN. Many of these issues later resurfaced in President Clinton's executive order.

Harry Harding suggests that Chinese concessions in 1990 were motivated by the threat of losing MFN and a desire to ease other U.S. sanctions, particularly restrictions on access to international loans. Chinese measures in 1990 included the release of some 800 dissidents, an agreement to purchase $2 billion of Boeing jetliners (with an option for additional purchases), the visit of a high-level trade mission to purchase U.S. goods, and allowing Fang Lizhi, a leader of the democracy movement who had taken refuge at the American embassy, to leave the country (Harding 1992, 260–65). In 1991 Chinese officials announced that they would refuse to accept conditional renewal of MFN, but made a series of concessions such as promising to ban the export of goods produced with prison labor, stopping illegal transshipments of textiles through third countries, pledging better protection for American intellectual property, and promising to purchase more American goods. Other conciliatory measures included placing a reactor sold to Algeria under IAEA safeguards, agreeing to discuss limits on Middle East arms sales, and considering joining the nuclear Non-Proliferation Treaty and adhering to Missile Technology Control Regime export restrictions. Chinese Premier Li Peng acknowledged in late June 1991 that China had made "a great deal of efforts" to preserve most-
favored-nation status (Harding 1992, 279).

President Clinton's May 1993 executive order imposed a series of specific conditions for the renewal of Chinese MFN status. There were two mandatory conditions: that MFN "substantially promote the freedom of emigration objectives" of the Jackson-Vanik amendment and that China comply with the terms of a 1992 bilateral agreement forbidding exports of products produced by prison labor. The executive order also called for "overall significant progress" toward:

- adhering to the Universal Declaration of Human Rights
- releasing Chinese citizens imprisoned or detained for the nonviolent expression of their political and religious beliefs, including such expression of beliefs in connection with the Democracy Wall and Tiananmen Square movements
- ensuring humane treatment of prisoners, by allowing international humanitarian and human-rights organizations access to prisons
- protecting Tibet's distinctive religious and cultural heritage
- permitting international radio and television broadcasts into China [This clause referred to Chinese jamming of Mandarin-language Voice of America broadcasts].

The executive order also called for efforts to ensure that China abided by its commitments to fair treatment for U.S. businesses, and adhered to the Nuclear Non-Proliferation Treaty, the Missile Technology Control Regime guidelines and parameters, and other nonproliferation commitments (Clinton 1993).

Assistant Secretary of State for East Asia Winston Lord testified before the Senate in March 1993, commenting on Secretary of State Christopher's visit to Beijing and assessing Chinese progress toward meeting the conditions in the executive order. Lord stated that the two mandatory requirements "seem to be in relatively sound condition." He noted that the Chinese foresaw no problems in dealing with the "less than ten" cases of families of dissidents who wish to leave China and had signed an agreement tightening the implementation of the 1992 bilateral agreement on prison labor and broadening the range of prisons that could be inspected by U.S. diplomats. Lord described modest progress on four of the five other objectives, but noted that "there was no progress during this trip on Tibet, with the single exception of their willingness to provide information on 106 Tibetan prisoners" (Lord 1994).
China also announced or instituted a variety of steps to tighten enforcement of previous agreements with the United States and to increase imports of U.S. goods. The Ministry of Trade announced that it had fined four firms caught evading U.S. textile quotas by exporting goods with false labels and banned them from exporting for one to two years (Reuters 1994b). Chinese officials issued a statement detailing intellectual property rights and claiming to handle 13,000 cases of copyright infringement and counterfeiting a year, including 500 involving foreign trademarks (Reuters 1994c). Chinese authorities announced a crackdown on compact disc piracy, including raids on 400 music stores and seizure of more than 53,000 pirated CDs, 2,700 pirated cassette tapes, and almost 100 fake laser discs (Associated Press 1994a). A Chinese trade and investment mission to the United States signed contracts worth over $11 billion, including at least $600 million worth of purchases of U.S. commodities (Reuters 1994d). Foreign Trade Minister Wu Yi stated at the signing that MFN was a trade issue and the basis of economic relations between the two nations. "Mixing it with other issues is not wise. We hope that a practical attitude will be adopted and man-made obstacles excluded, so that U.S.-China relations can develop in a stable and healthy way" (Reuters 1994e). China ordered nearly 800,000 tons of subsidized U.S. wheat under Washington's Export Enhancement Program, following a January purchase of 815,000 tons of U.S. wheat (Reuters 1994f). China also released Wang Juntao, a leader of the 1989 prodemocracy movement, from prison and allowed him to leave the country (Associated Press 1994b).

When the steps listed above are considered along with other Chinese moves such as signing the Nuclear Non-Proliferation Treaty and the Chemical Warfare Convention, agreeing to abide by the Missile Technology Control Regime, and acquiescing to U.S. actions in the Security Council, it is evident that China has made significant efforts to improve relations with the United States in the aftermath of Tiananmen Square. How much of that effort is related to pressure on MFN is another question. Many of the international actions are commitments that may or may not be carried out, depending on U.S. responses. China also has considerable leverage over the United States on the North Korean nuclear weapons issue. Chinese approval is necessary if the Security Council is to approve sanctions, and China must comply with the sanctions if
they are to be effective. This is another “indirect linkage,” but as Senator Sam Nunn noted, “Our top priority in dealing with China must be the situation on the Korean peninsula.” Nunn stated that “We’re concerned about human rights; that will continue to be our concern . . . but I think we need to find better tools to deal with it than the MFN treaty” (Reuters 1994a).

**CONCLUSION**

Analyzing China’s responses to U.S. threats of revoking MFN status by issue area presents a mixed picture. Chinese leaders have viewed economic and trade issues as a legitimate arena for discussions. MFN has been a relatively effective tool in winning concessions, although most of the concessions have actually come through bilateral negotiations and through the use of other policy tools such as Super 301. MFN has certainly influenced the timing of Chinese concessions, but concessions of a similar magnitude might well have been achieved with other policy instruments.

Foreign policy is harder to assess. China has certainly moved in directions favorable to the United States, but the linkage with MFN is unclear. China has strategic reasons for accommodating U.S. international interests that do not conflict with its own international objectives. Moreover, the extent of Chinese compliance with its international commitments is uncertain. MFN may have played a role in altering Chinese policy, but the linkage runs both ways—China’s critical role on the North Korea nuclear issue makes it difficult for the United States to withdraw MFN.

Human rights have enjoyed some successes. On the whole, however, the successes have been marginal improvements, which have not dramatically improved civil and political rights in China. In general, China has regarded U.S. efforts as violations of its sovereignty and unwarranted interference in its internal affairs. The regime has been unwilling to make any human-rights concessions, which might threaten its power, and the threat of MFN withdrawal is unlikely to alter its views.

The preceding analysis indicates that the increasing U.S. domestic and international costs of withdrawing China’s most-favored-nation status gradually eroded the credibility of this threat. Withdrawing MFN status would have hurt China more than the United States, but the absolute costs to the United States in terms of lost trade and investment opportunities came to dominate the political
debate. China was well aware of these costs and conscious that the United States did not want to carry out its threat. The mood in Congress gradually shifted from a desire to punish China to increasing concern over the impact on the U.S. economy, unemployment, and potential restrictions on access to Chinese markets. China held considerable leverage on international issues like North Korean nuclear weapons and felt that relatively minor political and economic concessions would be sufficient to win renewal.

Selective and carefully targeted sanctions are likely to be more effective than threats to revoke China’s MFN status. These kinds of sanctions do less damage to the U.S. economy, making them a more credible threat in negotiations. Most Chinese concessions have actually come in bilateral negotiations not explicitly linked to MFN. Selective sanctions can get results without the acrimonious public debates over renewal of MFN that strain Sino-U.S. relations and spill over into other areas of potential cooperation. A lower key approach that emphasizes diplomacy, bargaining, and face-saving measures will likely to be more effective in advancing U.S. interests than the confrontational tactics associated with MFN renewal. The Clinton administration’s decision to delink human rights and MFN status appears to recognize this reality.

Notes

1. This paper does not explicitly consider the impact of the Uruguay round tariff reductions, but the reader should note that the Uruguay round will lower MFN tariff rates substantially, increasing the spread between MFN and general tariff rates. The impact on China will be less than might be imagined, since tariff rates on the goods China imports to the United States were not dramatically reduced.


4. The figures here differ slightly from the World Bank figures, because the World Bank used an estimate of $16.261 billion for 1990 imports from China instead of $15.237 billion. The IBERC study does not report total decline in imports from China, but these figures can be calculated by multiplying the decline in trade by the volume of imports from China. These calculations correspond closely with those obtained by scaling the “Change in Chinese imports” to account for the total volume of trade.
China’s economy has consistently outperformed official growth targets in recent years. Meeting the official target of 9 percent actually implies a GDP growth rate 3 to 4 percent lower than private estimates.

It is also reasonable to assume that domestic distributors might absorb some of the costs of the higher tariffs rather than passing them on to consumers. This could be modeled by scaling up the costs to consumers by a factor lower than three.

The IBERC study is intended to emphasize the costs to U.S. consumers and does not present all the data needed to precisely calculate changes in domestic producer surplus and tariff revenue from other countries. Tariff revenues cited here do not appear in the IBERC study, but have been specifically calculated for this paper.

For example, China sold nuclear power plants to Iran, Pakistan, and Algeria shortly before signing the NPT; it has also been accused of selling M-11 missiles to Pakistan. A Chinese foreign-policy official recently admitted in private conversation with the author that China has sold “missile components” to Pakistan.

References

Agence France Presse. 1993. U.S.-China trade war would cost Hong Kong 70,000 jobs: Patten. May 8.

Associated Press. 1994a. April 20, 20:12:30 EDT.

Associated Press. 1994b. April 23, 15:36:08 EDT.


Hong Kong Economic and Trade Office. 1994. *Possible impact on HK if China lost MFN status* (March 7).


Reuters. 1994a. April 7, 13:33:19 EDT.


Reuters. 1994c. April 12, 14:07:05 EDT.

Reuters. 1994d. April 21, 10:46:50 EDT.

Reuters. 1994e. April 21, 19:40:07 EDT.

Reuters. 1994f. April 24, 20:44:07 EDT.


