Gulliver Untied: Entry Deterrence in a Unipolar World

Joanne Gowa∗
Department of Politics
130 Corwin Hall
Princeton University
Princeton, NJ 08544
jgowa@princeton.edu
609-258-5831

Kristopher W. Ramsay
Department of Politics
130 Corwin Hall
Princeton University
Princeton, NJ 08544
kramsay@princeton.edu
609-258-2960

September 14, 2011

Abstract

That the anarchic system generates incentives for states to balance each other’s power is conventional wisdom in international relations. As such, the contemporary unipolar system is an anomaly. The existing literature explains the persistence of a one super-power world in various ways, including U.S. benevolence; the constraints international institutions impose on U.S. power; the collective-action problem that would-be great powers confront; and the magnitude of U.S. military and economic resources endowment. We advance a new explanation for the stability of one super-power systems. Using a simple two-period model, we show a dominant power can exploit its first-mover advantage to partially sink the cost of war, thus making prospective challengers better off staying out of the great-power game. In contrast to other work, our explanation focuses on the super-power’s role in deterring balancing behavior and implies investment in larger standing military forces than might seem necessary to respond to existing threats.

∗We thank Brett Benson, Bob Keohane, and Andy Moravcsik for comments. As usual, all remaining errors are our own.
INTRODUCTION

That the anarchic structure of the international system endows on states incentives to balance each other’s power is conventional wisdom among students of international relations. The absence of a third-party mechanism to enforce the peace creates a self-help system in which the territorial integrity of each state depends only upon its own actions. A state can opt either to build up its own military strength or to join other states in an alliance strong enough to defend its members against an external threat. The idea that states in a self-help system have a stake in balancing each other’s power is among the “oldest and best known” in the field (Ikenberry, Mastanduno, and Wohlforth 2009, 18).

Because incentives to balance are a consequence of anarchy, they do not vary across international systems. Rather, as Kenneth N. Waltz observes, “balance-of-power politics prevail whenever two, and only two, requirements are met: that the order be anarchic and that it be populated by units wishing to survive” (1979, 121). As such, the contemporary international system is anomalous. Neither a single country nor any coalition of states has attempted to match U.S. power in the 20 years since the Soviet Union collapsed. This is so despite the fact that by “virtually any measure,” as Stephen M. Walt observes, the United States “enjoys an asymmetry of power unseen since the emergence of the modern state system” (2005, 31).

The existing distribution of power can aptly be described as unipolar—that is, it is a system in which one state is “unambiguously in a class by itself compared to all other states” (Ikenberry, Mastanduno, and Wohlforth 2009, 5). The margin of power the United States now enjoys allows it to act much more freely abroad than it could during the bipolar era, when it feared that intervening elsewhere might spark a Soviet reaction that could end up in a nuclear confrontation. As Robert Jervis notes, the most exigent issue at stake during bipolarity was whether a crisis would erupt that could “lead to superpower war.” The most
important issue now, he observes, is whether “unipolarity will last” (2008, 207).

The existing literature advances several explanations for the surprising durability of unipolarity. Among them are the constraints international institutions impose on the unilateral exercise of the unipole’s power; the collective-action problem that deters the formation of an effective countervailing coalition; and the prohibitive cost associated with challenging U.S. economic, military, and technological capabilities. In each case, agency is assigned to actors other than the unipole itself. Thus, for example, explanations based on institutions contend that several organizations of vital importance to the unipole wield a credible threat to expel it from membership should it attempt to extract excessive rents. In this view, the continued existence of the contemporary system depends on U.S. willingness to respect the constraints institutions impose on its use of power, making it the political analogue of a contestable monopoly.

Because extant explanations focus on exogenously imposed constraints, they do not attend to the possibility that the dominant state itself can act to secure its incumbency. We argue here that a first-mover advantage accrues to the incumbent unipole that can allow it to drive the net gains of would-be entrants to zero. In our argument, the unipole does not owe its privileged position to its respect for the constraints other actors impose on it or to the inability of prospective challengers to engage successfully in individual or collective balancing efforts. Instead, its continued existence is due to its nonmyopic response to the entry problem that makes challenges to its power costly on net.

We begin with a brief review of existing explanations of the stability of the contemporary system. Then, we analyze a sequential strategic interaction in which the dominant state has an opportunity to usurp the balancing dynamics that students of international relations have long believed inhere in anarchy. We show that the unipole can retain its singular position under some although not all possible configurations of the parameters. Because the contemporary system is the first instance of unipolarity since the advent of the state system
in 1648, we cannot systematically test our argument against the alternatives. We show, however, that it is the only argument that is consistent with the steady expansion of the gap between U.S. military expenditures and force deployments and those of other states after the end of the Cold War. Even if entry occurs, the incumbent can still use its first-mover advantage to determine the equilibrium level of arms that will result.

**EXISTING EXPLANATIONS**

Some observers argue that the absence of balancing is a result of the fact that the country that currently occupies the role of the pivotal player is the United States. Essentially, they argue that anarchy alone is insufficient to precipitate balancing. Rather, “it matters who the unipole is, where it is located, and how it chooses to use its power” (Walt 2008, 120). In their view, the values, interests, and actions of the United States render it the international equivalent of a benevolent dictator. By sheltering many states under its security umbrella, the United States persuades them to accept “its role as legitimate,” even as it transforms anarchy into hierarchy (Lake 2009, 176).\(^1\) Because no threat to their integrity, territorial or otherwise, exists, states have no reason to expend resources on balancing the power of the United States.

This argument echoes hegemonic-stability theory. According to it, a hegemon provides public goods to the international system, including the “security and protection of property rights” and a “liberal international economic order” (Gilpin 1981, 145). The hegemon’s large stake in maintaining a stable world order means that it is better off if it provides the public goods.\(^1\)

\(^1\)We do not discuss soft power here, as we agree with Jervis that “the distribution of most forms of soft power will roughly correlate with the distribution of economic and military resources” (2008, 192) and with the point made by Wohlfarth and Brooks (2005), who note that the empirical evidence is inconsistent with the claim that U.S. power is the target of “constraint” actions by other states (2005, 75).
goods necessary to do so even if it absorbs all the costs involved. The current system is
the international analogue of Mancur Olsons “privileged” group—that is, a group in which
the welfare-maximizing strategy of one of its members dictates supplying public goods if the
alternative is their absence, even if doing so enables widespread free riding (1975).

John A. C. Conybeare (1984) largely undermined the economic variant of hegemonic
stability theory when he pointed out that free trade is not the optimal policy of a single
large state. Given its market power, an optimal tariff is instead its welfare-maximizing
choice—that is, a tariff set such that the government revenue it generates exceeds the sum
of the deadweight loss and the loss in consumer surplus trade barriers produce. Free trade
maximizes the welfare only of small states—in particular, those states unable to manipulate
their terms of trade. The theory took another hit when critics pointed out that a hegemon
could opt to extract monopoly rents rather than supply public goods.2 John Gerard Ruggie
made this point forcefully when he asked whether Nazi Germany, had it emerged the victor
in World War II, would have behaved as the United States did (1982).

For some students of international relations, the appeal of a benevolent hegemon is less
a function of the public goods it provides than of its relatively modest aims. Josef Joffee
concedes that the United States “irks and domineers” because it tries “to call the shots
and bend the rules” (1997, 16). Walt, too, acknowledges that “Europeans may dislike U.S.
policies, Asians may worry about U.S. judgments, and Chinese leaders may see the United
States as a rival over the longer term” (2009, 103). Nonetheless, both Joffee and Walt
believe that U.S. actions, while sometimes annoying, are nonetheless unlikely to precipitate

2Some prominent analyses argue that a hegemon is doomed to shoot itself in the foot
because it “tends to overpay for security,” weakening “the international foundations of its
external security” (Layne 1999, 34). Robert Gilpin advances several reasons that the hege-
mon will inevitably decline (1981, chapt. 4). Common to these arguments is the belief that
the hegemon is destine to expand beyond the point at which the marginal cost of doing so
exceeds the marginal gain or to consume rather than invest. The supporting logic is weak,
however. See Rogowski, e.g., for a detailed explanation (1983,13–38).
balancing behavior because it is common knowledge that the United States does not “go to war for land and glory” (Joffee 1997, 16). They and other observers maintain that balancing is not inevitable but “cranks” up as a function of whether the dominant state behaves as an “usually placid elephant or an aggressive T. rex” (Joffee 1997, 16).

The uniquely American character of the current order plays a less important role in other explanations of unipolar persistence. As Walt notes, some observers disagree with the premise of the benevolent unipole explanation, maintaining that the United States is not a “status quo power” (2005, 25). Others claim that any highly skewed distribution of power represents a threat to the survival of states. Thus, for example, Christopher Layne observes that states do not worry about the hegemon’s intentions but about its power (1993, 4). Timothy Garton Ash concurs, noting that the real problem is “simply the power. It would be dangerous even for an archangel to wield so much power” (2005, 99).

Skeptics of the benevolent hegemon explanation prefer to emphasize the trigger mechanism international institutions have at their disposal. In their view, postwar institutions originated to deter the use of Soviet power but endure because they endow their members with the power to tie the unipole’s hands. As long as the incumbent unipole regards the integrity of existing institutions as vital to its national security, their less powerful members can constrain the unipole’s foreign policy generally and its trade policy in particular (Ikenberry 2003, 4). In doing so, they “legitimate America’s authority over others” (Lake 2009, 188).

Historically we have not seen the trigger mechanism activated because the United States has an interest in maintaining the integrity of existing institutions. Thus, for example, its membership in the World Trade Organization (WTO) deters it from arbitrarily increasing its trade barriers and has induced it to reverse policies that the organization has declared conflict with its rules. The leverage the organization exerts inheres in the fact that the trade-policy incentives of economically powerful states conform to those of the well-known
repeated Prisoner’s Dilemma (PD) game: each of them has an incentive to retaliate against a deviation by any other, making defection costly on net. Thus, playing by the rules of the game is the optimal strategy even for a state that stands alone in terms of its military power.

No analogous mechanism exists in the domain of security, however. Given the skewed distribution of military power, the dominant state is unlikely to face a credible threat strong enough to deter its unilateral exercise of coercion. That the United States went to war against Iraq in 2003 despite the refusal of United Nations Security Council to endorse its action attests to this. In a classic example of a collective-action problem, neither the UN nor any of its members either individually or as a coalition wielded a credible threat to retaliate: the costs of doing so were very high and the benefits were nonexcludable and nonrival in consumption. Although an effective institutional trigger mechanism may exist in the trade arena, no such credible threat constrains the U.S. pursuit of security goals.

Prospects of successful deterrence improve, of course, if a single state acquires enough power to punish hegemonic transgressions, making its threat to do so rational and credible—that is, punishing a deviation by the hegemon makes it better off than does overlooking it. In this case, it is possible to deter the existing unipole from wandering off the peaceful equilibrium path even in the security arena. During the bipolar era, for example, the Soviet Union wielded a highly credible threat to punish U.S. incursions into its sphere of influence. Indeed, absent its 1991 collapse, it is hard to imagine that the United States ever would have entertained the idea of stationing missiles in Poland or in the Czech Republic or of waging war against Iraq.

Some argue that unprecedented U.S. primacy, not institutions or benevolence, explains unipolar stability. A thorough accounting of power enables Wohlforth (1999) and Brooks and Wohlforth (2003) show the United States is the first leading state in modern history to hold a decisive edge in all underlying dimensions of power. Its economic power, the size of its military arsenal and the state of military technology, human capital, and geopolitical
location give the United States a decisive global advantage. In fact, Wohlforth observes, “[t]he raw power advantage of the United States means that an important source of conflict in previous systems is absent: hegemonic rivalry... No other major power is in a position to follow *any policy* [emphasis added] that depends for its success on prevailing against the United States in war or an extended rivalry (1999, 7).

While intuitively appealing, this argument does not explain the existing stock of U.S. arms and equipment. In both twentieth-century world wars, for example, the United States was also a major power with latent capabilities and in each case rapid ramping up of U.S. military forces provided the allies with a decisive margin of victory. In December 1941, the U.S. Army had 1.7 million men under its command; less than four years later, the number had jumped to 8 million. Army ground forces expanded during the same time period from 870,000 to 2.8 million.⁴

It is important if perhaps obvious to note that the post-Cold War U.S. military arsenal did not arise in a fit of absent-mindedness. As it did in the wake of both twentieth-century world wars, Washington could have opted to skew consumption toward butter rather than guns in the wake of the Soviet collapse. It could have demobilized, slowed its procurement of new weapons systems, closed overseas bases, and reduced research and development spending. But it made a very different choice, even after accounting for the costs of war in Iraq and Afghanistan and for the expenditures necessary to prosecute the war on terror. Even though no post-Cold War threat exists on a par with the Soviet threat during the bipolar era, the standing U.S. military capacity far exceeds that of any state that might conceivably challenge it.

The United States now maintain 837 military bases outside of the 50 states, more than 700 of them on foreign territory. The corresponding figure for China is eight and for Britain 28. Between 2001 and 2008, its defense budget rose by 70 percent net of the costs of war in

Table 1: Elements of Military Power 2009

<table>
<thead>
<tr>
<th></th>
<th>Overseas Bases&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Defense spending&lt;sup&gt;b&lt;/sup&gt; (2009 US $ m.)</th>
<th>Defense R&amp;D&lt;sup&gt;c&lt;/sup&gt; (2009 US $ m.)</th>
<th>Nuclear warheads&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>837</td>
<td>668,604</td>
<td>84,944</td>
<td>9,400</td>
</tr>
<tr>
<td>Russia</td>
<td>19</td>
<td>53,330</td>
<td>n.a.</td>
<td>12,000</td>
</tr>
<tr>
<td></td>
<td>[2]</td>
<td>[8]</td>
<td>[n.a.]</td>
<td>[127]</td>
</tr>
<tr>
<td>China</td>
<td>8</td>
<td>110,100</td>
<td>n.a.</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>[1]</td>
<td>[16]</td>
<td>[n.a.]</td>
<td>[3]</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>28</td>
<td>57,907</td>
<td>257</td>
<td>255</td>
</tr>
<tr>
<td></td>
<td>[3]</td>
<td>[9]</td>
<td>[3]</td>
<td>[3]</td>
</tr>
<tr>
<td>France</td>
<td>16</td>
<td>66,869</td>
<td>5,688</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>[2]</td>
<td>[10]</td>
<td>[7]</td>
<td>[3]</td>
</tr>
</tbody>
</table>


Afghanistan and Iraq. Even as it closed some bases, it has prepositioned military equipment around the world. Carrying tanks, munitions, food, water, and fuel U.S. ships at sea await orders to deploy to just about any point on the globe (Huff 2001, Chilcoat and Henderson 1994). No other country comes close to having a similar capacity to project power. Table 1 displays various components of U.S. military power relative to other large states.

Moreover, the gap between U.S. defense spending and that of its closest competitor has widened progressively since 1991. As Figure 1 shows, Russian military spending dropped sharply with the end of the Cold War. U.S. military expenditures, however, stayed remarkably constant, remaining at about the same level in real terms as they had in the 1970s and 1980s. As a result, a large gap has emerged between the United States and its closest competitor. This is so even though the U.S. expenditures that the figure displays are net of spending on the wars in both Iraq and Afghanistan. Moreover, the trend does not change markedly in the wake of the September 2001 attacks. In 2009, U.S. defense spending was more than double that of the next four largest states combined—that is, China, Russia, the
In the next section, we advance an explanation of this gap that assigns a central role to the U.S. effort to maintain its position as the incumbent unipole. We view the size of U.S. forces as a product of its effort to deter entry into the great-power arena. Our explanation does not invoke institutions, hegemonic benevolence, public goods, or collective-action problems. As no informational asymmetries exist, it also does not invoke reputation (cf. Alt, Calvert, and Humes 1988). Instead, we focus on the impact preemptive investment exerts on the equilibrium that prevails if entry occurs. As the next section will make clear, the argument

4In Alt, Calvert, and Hume (1988) the hegemon deters entry by fighting costly wars, establishing a reputations for high resolve.
we advance is a natural extension of Waltz’s seminal work (1979). In it he relied on what was a then state-of-the art analysis of interactions among firms in an oligopoly to capture the essence of balance-of-power politics in an anarchic international political system.\(^5\)

**ENTRY DETERRENCE**

In the dynamic situation we model here, the incumbent unipole can act preemptively to deter a challenge to its power. The dominant state is the political analogue of an incumbent monopolist, able to exploit its first-mover advantage to protect its monopoly. In the industrial-organization literature, the incumbent firm can invest in capital in the first period in order to set its own and the entrant’s second-period profits. Because its first-period investment can drive the entrant’s profits to zero or below, the existing firm can protect its position (Tirole 1989). In our model, the existing unipole has the option to sink resources into military capacity in the first period to lower its marginal cost of arming thereafter. This endows it with a credible threat to expand its military effort in the event entry occurs, raising the barriers to entry that face a prospective entrant into the great-power arena.

Although this argument is inconsistent with the long-standing belief that an anarchic system inevitably produces incentives to balance, it is nonetheless consistent with the view Pentagon officials expressed in their 1992 draft planning guidelines. The Cold War had barely ended when they made clear their belief that U.S. security required “deterring potential competitors from even aspiring to a larger regional or global role.”\(^6\) President George

\(^5\)The argument closest to ours is made in Posen (2003). HE contends that what he labels the *U.S. command of the commons*—that is, its acquisition of a ”military advantage at sea, in the air, and in space (2003, 21)–makes it impossible for any other state to enter the great-power arena. In terms of our argument described below, Posen argues that command of the commons has increased the fixed costs of becoming a peer-competitor of the United States, but does not make an argument about variable costs and commitment.

\(^6\)http://www.pbs.org/wgbh/pages/frontline/shows/iraq/etc/wolf.html accessed August
W. Bush reiterated this view in his speech to West Point cadets in 2002. “America has, and intends to keep,” he stated, “military strength beyond challenge, thereby making destabilizing arms races pointless” (cited in Ikenberry 2011, 256-57). The construction of a military force immune to any challenge emerges endogenously in the model we analyze here. In no other explanation of unipolar persistence is this true.

We model the interactions between an incumbent great power and a potential challenger. In our model, as in the real world, the payoffs of each state vary with its level of arms and that of the other state. All payoffs are common knowledge—that is, each state has complete information about its own preferences and payoffs and those of the other state. We consider the simplest dynamic problem: a two-period game. In it, the incumbent unipole acts nonmyopically—that is, it takes into account the impact its first-period investment exerts on its own and an entrant’s second-period cost of arming.

In the first period, the incumbent makes an investment in military capacity. In the second period, it and the entrant make two simultaneous decisions. First, each decides whether to participate in the great-power contest. Staying out yields a fixed payoff. If only one country participates in the second period, it selects its utility-maximizing level of arms given its opponent’s choice of a zero level of arms. If both opt in, they simultaneously choose their arms levels. We denote these as $a_i$ for the incumbent and $a_c$ for the challenger.

As is standard, the utility that a country accrues from arming depends on its own arms and on the arms of its opponent. We capture the benefit of participating in the great-power contest with a simple utility function that we label the “arms value” function—that is,

$$v(a_i - a_j).$$

Thus, the utility of a state, $v$, increases with the difference between its own level of arms and
that of its opponent. In terms of the calculus of our model, this means that the first derivative of the arms value function is positive ($v' > 0$)–that is, the change in utility associated with an increase in one’s own arms given a fixed level of the other’s arms is positive.

We assume that the marginal return to arms acquisition is decreasing when the opponent’s arms level is fixed. Diminishing marginal returns means that the utility a state derives from an additional increment of arms, holding constant the other state’s arms level, is smaller than is the utility it received from the preceding unit it acquired. Thus, the second derivative of the arms value function is negative ($v'' < 0$). A country’s utility increases in its arms level but at a decreasing rate.

The military arsenals of each country are complements–that is, the utility a state derives from its arms increases with the arms level of the other country. An often-cited example of complementary goods involves left and right shoes. The utility a left shoe offers a consumer increases in the presence of a right shoe and vice versa. In the great-power contest, arms are complements because the value to a state of acquiring more arms rises with the amount of arms the opponent accumulates. Thus, the value to a great power of building one more battleship is higher if its opponent already has two battleships than if it only has one.

Acquiring arms involves three types of costs: fixed, capacity, and marginal costs. As in any other industry, arms acquisition requires expenditures on fixed costs–that is, the costs necessary to produce a given good at any positive level. These costs are independent of output levels. As they are dedicated assets, fixed costs are sunk–that is, they cannot be profitably redeployed to the production of other goods, rendering them irrelevant to future decision making. In the great-power game, fixed costs involve the expenditures necessary to acquire the infrastructure that is essential to using military force, including, for example, building command and control structures and educational and training facilities that enable future soldiers to design and implement military strategy. These and other expenditures create the foundation of a modern military and must be made prior to or at the same time
as hardware and personnel are acquired.

Variable costs involve both capacity and marginal costs. Capacity costs are those associated with the acquisition of the arms stocks that prosecuting war requires. They vary not only with the size of the stock but also with its allocation among different and to some extent substitutable goods—e.g., tanks, drones, soldiers, aircraft carriers. They also depend on the size of the research and development effort that ensures state-of-the-art weaponry. Finally, equipping battle-ready forces involves variable costs—that is, those involved in deploying troops to battle, including expenditures on labor, fuel, and on support and combat soldiers.

As the incumbent, the dominant state incurs its fixed and capacity costs in the first period—that is, before a prospective great-power rival decides whether to enter the game. It can also expend resources on both capacity and arms in the second period. As such, the total costs of the incumbent great power, which we denote by $c_i$, equal:

$$c_i = \begin{cases} 
    f_i + r_i k_i + m_i a_i & \text{if } a_i \leq k, \\
    f_i + (r_i + m_i) a_i & \text{if } a_i > k.
\end{cases}$$

$f_i$ is the fixed component of the incumbent’s cost. We denote the level of capacity as $k_i$, which is acquired at a per unit cost of $r_i$. The per unit cost of acquiring arms equals $m_i$. Thus, the total cost of capacity is $(r_i \times k_i)$ and the total cost of acquiring arms is $(m_i \times a_i)$. We assume that the incumbent’s costs are all positive.

The incumbent’s cost function shifts between these equations as the relationship between $a_i$ and $k_i$ changes. The first equation expresses the incumbent’s cost when it chooses a second-period level of arms that is no greater than its first-period capacity. This means that it pays only the cost of arming in the second period. If the incumbent chooses an arms level in excess of its first-period capacity, however, its second-period costs rise, as it must
acquire capacity as well as arms. The second equation expresses this case. It shows that the incumbent’s marginal per unit cost of acquiring arms rises to \((r_i + m_i)\). It is because the incumbent’s first-period investment decision affects its second-period cost that its ability to act preemptively influences the entry decision of the prospective challenger.

The challenger can only invest in the second period, making its cost structure simpler. Its marginal cost of arming includes both capacity costs and those associated with labor, ammunition, and fuel, i.e., \((r_c + m_c)\). The following equation, therefore, expresses its cost:

\[
c_c = f_c + (r_c + m_c)a_c.
\]

Having described the value and cost functions of each state, we can now begin to consider the requirements of equilibrium. We apply the concept of subgame perfection, which is a Nash equilibrium refinement that requires a player to choose the welfare-maximizing alternative whenever it acts. That is, a subgame perfect Nash equilibrium requires each actor to play a Nash equilibrium in every subgame. This means that only credible threats affect decision making. As reasoning backward from the second stage makes it easy to understand the strategic incentives and actions of both states, we first consider second-period actions.

As rational utility maximizers, each player takes the arms level of its prospective opponent as fixed and the prospective entrant treats the incumbent’s level of capacity as a parameter. This enables us to identify the optimal arming response. A best-response function describes each player’s optimal choice of arms for each possible level of its opponent’s arms—that is, it describes the set of points at which the marginal return to an additional unit of arms equals the marginal cost of its acquisition given an opponent’s arms choice. The level of military arms the incumbent chooses, therefore, solves one of the three following equations
Figure 2: Marginal cost and benefit for arming at level $a_i$, given a $a_c$ fixed.

that maximizes its expected utility:

$$v'(a_i - a_j) - m_i = 0,$$
$$v'(a_i - a_j) - (r_i + m_i) = 0,$$
$$k = a_i.$$

This differs from the single first-order condition that determines the challenger’s optimal response because the incumbent’s piecewise utility function reflects its ability to make a pre-contest investment. If the incumbent chooses to arm in the second period at a level less than its first-period investment of $k$, its marginal cost of arming is $m_i$. If it chooses a level that exceeds $k$, its marginal cost rises to $(r_i + m_i)$. When its choice equals $k$, its utility function is not differentiable.

Figure 2 illustrates its optimal choice in the second period. The downward sloping curves
trace out the marginal benefit that accrues to the incumbent from choosing different levels of arms \((a_i)\) given a fixed level of challenger arms \((a_c)\). If the incumbent is choosing arms at a level less than \(k\) then the marginal cost is \(m_i\), if the incumbent is choosing arms at a level great than \(k\) then the marginal cost is \(r_i + m_i\), and at \(k\) we can see that the incumbent’s utility function is not differentiable. As the figure shows, if we change the challenger’s level of arms to \(a'_c > a_c\) we see that the incumbent’s marginal benefit curve for arming shifts to the right, increasing its optimal level. Also, if the optimal arms level does not equal \(k\), the implicit function theorem can be used to sign the partial derivative of the incumbent’s optimal level of arms with respect to a change in the challenger’s level of arms--that is,

\[
\frac{\partial a_i^*}{\partial a_c} = \frac{v''(a_i^* - a_c)}{v''(a_i^* - a_c)} = 1 > 0. \tag{1}
\]

Because there are diminishing marginal returns to arming, \(v''(a_i^*, a_j)\) is negative and the partial derivative in equation (1) is positive. As such, arms are “strategic complements”: the incumbents best response to an increase in the challenger’s level of arms is to increase its own.

We model the incumbent unipole’s advantage by assuming a challenger can arm only in the second period. As before, a rational actor sets its arms level such that the marginal benefit of an additional unit equals its marginal cost--that is, it solves this equation:

\[
v'(a_c - a_i) - (r_c + m_c) = 0
\]

In a fashion paralleling the analysis for the incumbent, arms are also strategic complements here, so that increases in the incumbent’s arms level increase the prospective challenger’s optimal arms level.

Figure 3 (a) illustrates the best-response functions for both the incumbent and the chal-
lenger. Each slopes upward, reflecting the strategic complementarity between their arms. The function labeled $\bar{B}_i(a_c)$ is the best-response function of an incumbent power that did not to sink any costs into purchasing capacity in the first period. If it decides to play in the second period, therefore, it must expend costs on both capacity and arms—that is, $(r_i + m_i)$. The function labeled $B_i(a_c)$ is the incumbent’s best response when it pays only $m_i$, the marginal or variable cost associated with acquiring arms. This function traces out the best response of an incumbent that has invested in capacity in the first period. Finally, the function labeled $B_c(a_i)$ is the best response of the challenger. The challenger has only one such function because it cannot invest in the first period, making its marginal cost always equal to $(r_c + m_c)$.

Focusing on a few points in panel (a) of Figure 3 allows us to describe great-power interactions under various conditions. The intercepts $(\bar{H}_i, 0)$ and $(H_i, 0)$ indicate the level of arms that each player would choose if the incumbent were the only participant in the second-period contest. If the incumbent does not play, $a_i$ equals 0. Then, the optimal level of arms for the challenger is $H_c$, the level at which its best-response function crosses the
$a_i^*$ with entry $\bar a_i^*$ with entry $H_i$ without entry $\bar a_i^*$ with entry $\bar a_i^*$ with entry $k$ with entry $\bar H_i$ without entry $\bar H_i$ without entry

Figure 4: Effect of investment level $k$ on incumbent contest strategy

If it is the challenger that stays out, then the optimal choice of arms for an incumbent that has a high marginal cost of arming—i.e. one that must pay both capacity and material costs—equals $\bar a_i^*$. If, on the other hand, the prospective entrant stays out, and the incumbent’s marginal cost of arming is low, it will choose a higher level of arms $H_i$.

Figure 3 also displays the optimal choices of each state when both participate in the great-power contest. If the incumbent has to acquire capacity as well as arms in the second period, the equilibrium allocation occurs at $(\bar a_i^*, \bar a_c^*)$; otherwise it occurs at $(a_i^*, a_c^*)$. This makes clear why the first-mover advantage of the incumbent unipole matters. An incumbent that sinks cost into capacity in the first period reduces its marginal cost of second-period arming, shifting its second-period best response curve to the right. This means that the intersection of the players’ best-response curves occurs at higher levels of arms for each state than would be the case otherwise—that is, a first-mover advantage enables the unipole to force a welfare-maximizing entrant to acquire more arms than would otherwise be optimal for it. As we show below, raising the cost of entry can drive the net utility associated with it to zero, or below, for some prospective challengers.

Next, we focus on the investment decision itself. We begin by first setting first-period capacity exogenously—i.e., we fix $k_i$ at different levels—in order to make the dynamics clear. As seen in Figure 4, we can fix a level of investment $k_i$ at a point on the $a_i$ axis. The effect of an investment level $k$ changes as the it passes through different cutpoint values. For example, if $k < \bar H_i^*$ then the investment has no effect on the equilibrium level of arms

---

18
whether a challenger enters or not. But if $a_i^* < k < H_i$, then the optimal choice by the incumbent is $k$ whether the challenger enters or when it is deterred and remains out.

Panel (b) of Figure 3 gives an example of a $k$ between $a_i^*$ and $a_i^*$ where the best response of the incumbent country first follows that of an incumbent power with low marginal costs of arming, then progresses to behaving as an incumbent with high marginal costs of arming. The bold line denoted $B^k_i(a_c)$ is the incumbent’s complete best response for this $k$, and by choosing $k$ the incumbent power can essentially choose the equilibrium allocations of arms on the bold line connecting $(a_i^*, a_c^*)$ and $(a_i^*, a_c^*)$.

This resembles the industrial-organization case in which an existing monopolist chooses its optimal output on the best-response curve of a potential entrant such that staying out becomes the latter’s welfare-maximizing choice. The model we analyze here differs in that the unipole cannot actually set optimal second-period arms levels ex ante. Its preemptive investment, however, does allow it to affect the equilibrium level of second-period arms: sinking some costs of arming in the first period allows it to wield a credible threat to arm more heavily in the second period that it would otherwise possess.

Now, we explicitly describe the second-period entry decision. For simplicity, assume that the unipole is always better off participating, while the entrant can be indifferent between entering and staying out. Figure 5 adds a set of curves that trace out the arms allocations for each state that make the prospective entrant’s utility the same whether it challenges or accepts the status quo. The indifference curves represent different levels of utility. As before the challenger’s best response function, conditional on entering the contest, is $B_c(a_i)$. The bold line, $B^k_i(a_c)$, is the incumbents best-response curve, given its investment in capacity and the challengers level of arms.

As in Figure 3, $(a_i^*, a_c^*)$ are equilibrium levels of arms in the game between a “high cost” incumbent and a challenger; $(a_i^*, a_c^*)$ are the equilibrium arms levels between a challenger and a “low cost” incumbent. The vertical curves denoted by $U_c$, $U_{c'}$, and $U_{c''}$ are the challenger’s
indifference curves. These curves trace out combinations of arms, $a_i$ and $a_c$, that give the challenger a utility equal to the utility of non-entry, $u_c$. The location of the curve varies with the precise nature of the contest value function and of the fixed costs associated with arming, capacity costs, and labor costs. Finally, as the utility of the challenger is concave, “better than” sets are convex and the gradient vector is

$$\nabla U_c(a_c, a_i) = \begin{bmatrix} v'(a_c - a_i) - (r_c + m_c) \\ -v'(a_c - a_i) \end{bmatrix}.$$  

Because the term in the gradient related to the incumbent arms level, $a_i$, is negative, allocations to the right are worse than non-entry for the challenger and allocations to the left are better irrespective of the location of any particular indifference curve.

The central idea of our argument is that an incumbent great power can invest preemptively in its military in order to deter potential challengers. A comparison of indifference
curves and equilibrium arms levels allows us to predict when an incumbent great power will sink costs in military capacity in order to credibly commit to a level of arms in the second-period that is higher than it would be otherwise and when it can achieve deterrence only by acquiring excess military capacity—that is, capacity that will go unused in equilibrium. An equilibrium in which capacity is not actually used in equilibrium is the outcome that we believe best approximates that of the current system.

To begin the analysis consider a situation in which the cost of the military challenge generates a level curve like $U_c$. The curve’s location implies that the incumbent unipole does not need to invest in first-period capacity to make the challenger better off staying out than entering. This outcome is closely related to what economists describe as blockaded entry—a situation where a monopolist’s equilibrium production schedule absent any pre-competition investment creates a situation in which no other firm is better off entering than staying out. In the great-power contest, a similar outcome prevails when the comparative advantage of the incumbent power in the production of military capabilities alone suffices to keep the potential challenger out. This special case of blockaded entry is consistent with Wohlforth’s (1999) argument regarding U.S. dominance. Below we consider other cases observationally equivalent to this argument when considering the stability of unipolarity, but describe other aspects of the empirical environment like excess investment in military capacity.

Next, consider the other extreme. Suppose that the parameters of game imply that a military challenge to the incumbent generates the level curve $U_c''$. In this case, entry makes the prospective challenger better off than staying out for a wider range of arms levels. As a result, the incumbent unipole cannot deter entry but can only set its level of military capacity $k$ at the point between $(\bar{a}_i^*, \bar{a}_c^*)$ and $(\bar{a}_i^*, a_c^*)$ that maximizes its contest payoff. Thus, while it must accommodate entry, the incumbent can still determine the conditions under which it occurs.

The last situation we consider is the one in which the incumbent power chooses an
investment strategy that deters entry. Here the parameters of the great-power contest place
the challenger’s non-entry indifference curve between the equilibrium allocations of the high-
cost and low-cost incumbents in the contest. This produces either a full-commitment case or
a case of excess capacity. Suppose that \( \bar{a}_i^* < H_i \) and the critical indifference curve intersects
the challenger’s best-response curve at a point where the \( a_i \) coordinate is also less than \( H_i \),
as seen in Figure 6 (a). We call this the full-commitment case. Let \( \hat{a}_i \) be the first coordinate
of the point where the challengers indifference curve intersects his best-response curve. In
this instance, if

\[
v(\hat{a}_i) - (r_i + m_i)\hat{a}_i \geq v(\bar{a}_i^* - \bar{c}_i^*) - (r_i + m_i)\bar{a}_i^*
\]

the incumbent great power will invest \( k = \hat{a}_i \) and the challenger will stay out. This is a
case in which the incumbent’s first-mover advantage allows it to deter entry. Otherwise,
the incumbent will not invest, the challenger will enter, and the contest will be played with
equilibrium armaments \((\bar{a}_i^*, \bar{a}_c^*)\). In any circumstance in which the critical indifference curve
intersects the challenger’s best response at an \( \bar{a}_i^* < a_i < H_i \) the incumbent’s capacity invest-

Figure 6: Two cases of deterrent investment in military capital
ment generates a credible commitment to arm at a level $k$, which makes entry unattractive for the challenger. In this subcase, any capacity purchased in the first period is used and the incumbent’s equilibrium arms choice is $a_i^* = k$.

In the second subcase, seen in Figure 6 (b), a deterrent strategy generates a level of first-period investment such that excess capacity results—that is, capacity that the incumbent actually does not use in the second period. To see this, suppose that the parameters of the great-power contest imply that the challenger’s critical indifference curve intersects its best-response function at a point where its first coordinate is between $H_i$, and $a_i^*$. An example of such a situation is one where the critical indifference curve of the challenger is $\tilde{U}_c$ in the figure. In this case, the incumbent has an incentive to make a strategic investment that deters entry by the challenger. In the next period, its equilibrium level of arms is less than its military capacity.

In particular, it is optimal for the incumbent to choose $k'$ such that $(k', a_c^*(k'))$ is the contest allocation that makes the challenger indifferent between staying out of the contest and entering and contesting at $(k', a_c^*(k'))$. This is the excess-capacity case because $k'$ is greater than the amount of arms the incumbent chooses when the challenger fails to enter, $H_i$. That is, the incumbent invests at $k'$ to deter a challenge but will only utilize an amount of arms $H_i < k'$; thus, the incumbent holds but does not use excess capacity in order to deter entry. It is interesting to note that the challenger would enter the great-power contest if the incumbent chose the arms level it did when the challenger stays out ($H_i$) or had the incumbent chosen any capacity less than $k'$ in the first period. Thus the commitment power of excess capacity is key to entry deterrence in this final case.

These results are robust to situations in which more than one potential entrant exists or the game extends beyond two periods. Suppose multiple potential challengers exist. If the incumbent’s action deters the strongest among them, it will also deter every other challenger. Moreover, it doesn’t matter whether the challenger is only one state or a coalition of states.
It is also easy to see that extending the sequence of interactions does not affect the outcome. If the existing unipole faces repeated decisions about capacity investment, its incentives to build capacity and choose an arms level are the same as those that exist in the two-period model. Thus, the argument continues to hold for strategic interactions that extend beyond two periods. As such, the logic of the model is consistent with a variety of situations that can exist in a unipolar system.

**CONCLUSION**

The argument we present here resembles that of analyses of the impact of capital investment on entry in the industrial-organization literature (e.g., Dixit 1979; Fudenberg and Tirole 1983). In general, the difference is that the security dilemma makes arms strategic complements and market competition generally makes output strategic substitutes. Because reaction curves slope upward, equilibrium deterrence in international relations can generate excess military capacity. In contrast, the negative slope of the relevant curves slope in market competition implies an absence of excess capacity in equilibrium (Dixit 1979).

Our explanation of unipolar persistence assigns agency to the unipole, which can exploit its first-mover advantage to maintain its dominance. Whether balancing occurs is a function of the parameters of a strategic interaction that the incumbent’s own action influences. As such, the durability of a unipolar system does not depend on whether the dominant power is a benevolent or status-quo power; a member in good standing of various institutions; or on whether the forces in being it maintains are larger than those of other states. Instead, the critical issue is whether its ability to act preemptively enables it to make entry costly on net. Our analysis, extending Waltz’s to a case in which dynamics matter, demonstrates that one of his principal conclusions does not necessarily follow: anarchic international systems do not always create incentives for balancing.
Although it may be premature to claim that our argument explains continuing U.S. dominance, we note that it is consistent with the secular increase in the size of the U.S. arsenal and its continuing modernization in the absence of an imminent threat. Neither the alleged benevolence of the United States nor its implicit pledge to maintain the status quo can explain this trend. Nor is it explicable in terms of the constraints institutions impose or of the collective-action problem confronting would-be entrants. Our interpretation of the contemporary system implies that the dominant power itself may be indefinitely able to preempt a process that has long been understood as central to state survival under anarchy.
REFERENCES


