ARTICLES

JPART 21:i203–i232

PUTTING STREET-LEVEL ORGANIZATIONS FIRST: NEW DIRECTIONS FOR SOCIAL POLICY AND MANAGEMENT RESEARCH

The Organization of Discipline: From Performance Management to Perversity and Punishment

Joe Soss*, Richard Fording†, Sanford F. Schram‡
**University of Minnesota; †University of Kentucky; ‡Bryn Mawr College*

ABSTRACT

Drawing on participant observation, in-depth interviews, and statistical analysis of administrative data, this article explores the operation of performance management in the Florida Welfare Transition program and its effects on decisions to sanction welfare clients. Unlike most econometric research on welfare sanctions, we approach sanctioning as an organized practice that reflects, not just client characteristics and behaviors, but also organizational needs, routines, values, authority relations, environments, and systems of reward and punishment. Our analysis focuses on the organization of discipline and, in the process, suggests that scholars may misrepresent and misinterpret the incidence of discipline when they fail to account for the dynamic ways that organization and management shape sanctioning patterns.

INTRODUCTION

Over the past few decades, poverty governance in the United States has been transformed by the convergence of two powerful reform movements. The first, often referred to as "paternalist," has shifted welfare provision from an emphasis on rights and opportunities to a stance that is more directive and supervisory in promoting preferred behaviors among the poor. The second, often described as "neoliberal," has shifted governance away from federal government control toward a system that emphasizes policy devolution, privatization, and performance competition. Thus, in the era of neoliberal paternalism, lower level actors and private providers have been given greater policy discretion and have been called on to use their discretion in ways that enforce obligations and curtail deviance among the poor. Poverty governance has become more dispersed in its organization, more muscular in its normative enforcement, and more firmly rooted in the market logics of performance, profitability, and competition.

The resulting system entails a complex blend of autonomy and discipline. As decision-making responsibilities have become more dispersed, new incentives, penalties, routines, and management systems have been designed to encourage particular ways of thinking about

We thank Sarah Martyn Crowell for research assistance, and we are grateful to representatives of the Florida Department of Children and Families and Workforce Florida, Inc., for the data used in this article. Address correspondence to the author at jbsoss@umn.edu.

doi:10.1093/jopart/muq095

choice and, ultimately, to raise the odds that preferred paths will be taken. Welfare reform is widely viewed as an effort to redirect client behavior, but it is also an effort to discipline thought and behavior in service-providing organizations. In what follows, we show how such efforts are plagued by contradictions, perversities, and failures, yet, at the same time, ''succeed'' in shaping service provision so that street-level bureaucrats use their disciplinary powers to enforce work and penalize the most vulnerable segments of the poor.

Under the Temporary Assistance for Needy Families (TANF) program, welfare operations have been reorganized to reflect the principles of ''the new public management (NPM)''—a reform movement that has sought ''to replace traditional rule-based, authority-driven processes with market-based, competition-driven tactics'' (Kettl 2005, 3). Decentralization is a guiding principle of the new regime. Thus, policy authority has been devolved to facilitate locally tailored problem solving (Gainsborough 2003); contracts with private providers have been used to harness market incentives for efficiency and innovation (Dias and Maynard-Moody 2006; Heinrich 2000); and case managers have been given broader mandates to assess needs and allocate benefits, services, and penalties (Brodkin 1997, 2011; Hasenfeld et al. 2004).

Yet as lower level discretion has expanded, so too have efforts to discipline the *use* of discretion by structuring incentives and routines and enhancing pressures to ''perform'' in meeting program goals. ''The beginning of the twenty-first century,'' Moynihan (2008, 3) writes, ''finds us in an era of governance by performance management.'' In welfare-to-work programs, performance systems now serve as the core technology for monitoring operations and imposing accountability. They guide decisions about when to renew or terminate contracts with local providers; they provide state officials with a yardstick and a prod for the achievement of program goals; and they constitute the major way in which state TANF programs are evaluated by federal officials (Ewalt and Jennings 2004; Ridzi 2004). By establishing outcome benchmarks focused on work participation and placement, higher level officials define the goals of service provision and the terms of its evaluation (Brodkin 2011). Through sophisticated information systems, they monitor frontline activities and measure priority outcomes. And based on assessed performance, they use financial rewards and penalties to incent preferred behaviors and bring lagging service providers to heel. For local actors, choices are generally limited to the specific means that will be used to pursue mandated ends and are shaped by strong performance pressures and incentive structures.

Performance management functions, in all these respects, to discipline the ways that decentralized actors think about and make decisions. Yet proponents rarely conceptualize or analyze it in disciplinary terms. Performance management is typically presented as a way to harness the dynamic energies of markets, improve the evidentiary basis for policy choices, and reconcile policy experimentation with public accountability (Talbot 2005). The implicit promise is that local actors will be freed to go their own ways and then, later, will be judged by their performance and given the information they need to improve. The reality, however, involves a more complex interplay of structure and agency (Brodkin 2007, 2011; Moynihan 2008, Radin 2006). The focusing effects of outcome benchmarks, the pressures of competition, the prospects of incurring rewards or penalties, the awareness that one is being closely monitored: these features of performance management do more than just make agents accountable; they reshape agency itself.

In this article, we present an empirical critique of performance systems and the NPM. We show how the internal contradictions of NPM powerfully influence behavior at the frontlines of welfare reform but, at the same time, produce perverse organizational responses that subvert its underlying goals. Under the NPM, market logics, incentives, and penalties function to cultivate particular habits of mind (Foucault 1997), but they often do so in self-defeating and goal-displacing ways. The neoliberal script written by the NPM does not strictly control local service providers because it is, at root, a contradictory script that points local actors in multiple directions and provides resources for its own subversion. Thus, unlike many critiques of performance management, we do not treat perverse organizational responses as corruptions of or deviations from the NPM. Rather, we argue that they are predictable products of core contradictions within the NPM itself. Echoing Foucault (1980), we suggest that the disciplinary power of the NPM shapes consciousness and behavior in ways that are deep and far reaching yet also fractured, inconsistent, and incomplete.

Our second goal in this article is to clarify, from an organizational perspective, how efforts to discipline frontline service providers (through incentives, routines, pressures, etc.) shape efforts to discipline welfare recipients (through the use of benefit sanctions for noncompliance). Under paternalist welfare reform, human service organizations are called on to pursue a kind of "transformative moral work" that aims to reform aid recipients and move them into mainstream institutions (Hasenfeld 1992). Toward these ends, mandated client behaviors are closely monitored, incentivized, and made subject to penalties (Mead 1998, 2004). In this process, the paternalist tool of discipline *par excellence* is the sanction—a penalty that reduces or eliminates aid when clients fail to follow behavioral requirements. In a system where aid is conditioned on behavior, sanctions put muscle behind a host of participation requirements. They are key policy tools in the daily work of welfare case managers, and they have played a critical role in driving the major outcomes of welfare reform (Pavetti et al. 2004).

Researchers using econometric methods have produced a detailed portrait of how client characteristics relate to penalty patterns (Fording, Soss, and Schram 2007; Kalil, Seefeldt, and Wang 2002; Pavetti, Derr, and Hesketh 2003; Wu et al. 2006). To date, however, such studies have paid little attention to how sanctioning may be influenced by performance management or other aspects of organizational structure, process, and culture. The omission is striking because it ignores the obvious fact that sanction decisions are made in the context of organizational routines, by actors who occupy specific organizational positions. Leading field studies of welfare implementation reproduce this blind spot in reverse, revealing managerial and organizational dynamics while saying little about disciplinary practice (Lurie 2006; Riccucci 2005). As a result, scholars have failed to address one of the most distinctive and critical features of contemporary poverty governance: the interplay of systems for disciplining clients (e.g., sanctions) and systems for disciplining service providers (e.g., performance management).

In our second and third empirical sections, we seek to bridge this divide by presenting a mixed-methods study of how discipline operates on "both sides of the desk" at the frontlines of welfare reform. Using administrative data from the Florida Welfare Transition (WT) program, we establish a clear pattern linking performance pressures on providers to sanctions imposed on clients. Turning to field research on case manager discretion, we then explain this relationship and clarify its underlying mechanisms. Contrary to economist's account of rational "creaming," our field research suggests that this dynamic does not follow a simple logic of goal maximization. Rather, it is structured and fueled by the organizational conditions under which case managers operate. In this sense, our analysis underscores that sanctions are more than just responses to individual misbehavior. Because sanctioning practices must be organized, sanctioning patterns are, to a large extent, organizational products. To explain them, we must understand how the work of welfare provision is organized and managed, why it operates as it does, and how organizational structures, routines, and priorities shape disciplinary action.

CASE SELECTION AND RESEARCH METHODS

Our study focuses on the Florida WT program, not because it is typical of all state TANF programs but because it provides an ideal opportunity to analyze disciplinary action in a decentralized system of poverty governance. As a leading practitioner of "second-order devolution" (Gainsborough 2003), Florida has constructed one of the most decentralized TANF programs in the country. Frontline services have been contracted out to nonprofit and for-profit providers throughout the state, and primary authority over providers and local operating procedures has shifted down to 24 local public/private "Regional Workforce Boards" (RWBs). Under the oversight of a statewide public/private partnership called Workforce Florida, Inc. (WFI), Florida operates one of the most locally controlled and privatized TANF programs in the American states (Botsko, Snyder, and Leos-Urbel 2001, 7).

Florida also scores high on factors that raise the importance of sanctions. Under welfare reform, it adopted "some of the strictest time limits and work requirements in the nation" and broadened the pool of clients subject to sanctions by creating "few possibilities for exemptions" (Botsko, Snyder, and Leos-Urbel 2001, 4). The sanctions themselves impose a high level of penalty, resulting in an immediate, full-family loss of TANF benefits and a reduction of Food Stamp benefits to the fullest extent permitted by federal law (Botsko, Snyder, and Leos-Urbel 2001, 6). Moreover, Florida WT providers employ these sanctions at an extremely high rate compared to other states with full-family sanctions.[1] During the period of our study, sanctions were most commonly applied for violations of work-activity requirements (Agency for Workforce Innovation 2004) and were the most frequent cause of TANF case closings, routinely accounting for 40%–45% of closings (Florida Department of Children and Families 2009).

Our focus on Florida is also motivated by the state's heavy reliance on a competitive performance management system. Under this system, a state board negotiates with each RWB to establish a region-specific set of performance goals. Performance relative to these goals is then measured and used as a basis for state-level evaluations of regions as well as RWB evaluations of contracted service providers. Contract payments are tied directly to the achievement of performance goals, and local contracts often specify additional goals related to statewide performance measures.

Performance in the WT program is tracked on a monthly basis, and results are disseminated to providers and the public via "the red and green report"—so called because it ranks the 24 regions based on performance scores, coloring them red if they are in the

1 Applying the panel method advocated by Pavetti, Derr and Hesketh (2003) to Florida administrative data, we find that 47% of adults entering TANF in November 2001were sanctioned at least once in the next 18 months. Using a similar method, Pavetti et al. (2004) report that Illinois and New Jersey had full-family sanctioning rates of 13% and 17%, respectively, during the same period. Based on a period of 10 months, Pavetti et al. (2004) report a full-family sanctioning rate for South Carolina of 5%. The comparable sanctioning rate in Florida was 43%.

bottom quartile, green if they are in the top quartile, and white if they are in between. During the period of our study, the red and green reports included three items bearing specifically on the WT program: the "entered employment rate" among program leavers, the "employment wage rate" based on leavers' average initial wage, and the "welfare return rate" based on the percentage of clients who left for employment but later returned to WT.

Performance on the red and green report is taken very seriously at all levels of the WT program. Green scores can qualify a region for funding to undertake program improvements and allow a region to enter competitions for additional resources allocated by WFI. For service providers, red scores can mean the difference between contract renewal and termination. Between these extremes, providers typically lose "pay points" as a result of weak performance rankings and can draw unwanted scrutiny from the state board if their performance rankings fall.

As a result, the Florida WT program offers a superb opportunity to examine the convergence of strong forms of sanctioning, decentralization, and performance management. The program is not representative in the microcosmic sense of a statistical sample; it is representative in the analytic sense that it offers an unambiguous crystallization of policy developments that have occurred in varying shades of gray throughout the United States. To analyze the organization of discipline in this system, we adopt a mixed-methods approach that combines intensive, ongoing field research in purposively selected local one-stop centers with statistical analysis of statewide administrative data from the WT program. From March 2005 to January 2008, we conducted in-depth interviews with state-level officials, regional board members, program supervisors, and case managers. We observed sanction training workshops for case managers, meetings of region-level staff, and intake and orientation sessions with new WT applicants. Our field research focused most intensively on four of Florida's 24 workforce regions, where we conducted more than 60 in-depth interviews with all relevant frontline workers. Data for our statistical analyses are drawn from monthly administrative records for all participants in the WT program, January 2000 through March 2005. These individual-level data are supplemented with measures of community and organizational characteristics as well as measures of performance across regions and time.[2]

PERFORMANCE PRESSURES AND ORGANIZATIONAL BEHAVIOR: PERVERSITY IN THE FIELD

In principle, the NPM model suggests that local organizations will respond to performance pressures by pursuing innovations that advance statewide goals and improve the programs available to clients. Devolution should free local actors to experiment with diverse service delivery approaches in response to local needs. Statewide performance feedback should provide local actors with the data they need to learn from their own mistakes and achievements, identify successes in other regions, and emulate best practices. Competitive

2 To gain leverage on questions of validity in our field research, interviews were conducted separately by two of the three investigators, who then compared findings and interpretations to identify points of convergence as well as tensions that merited further investigation. All interviews were recorded and transcribed verbatim for analysis. Analysis of the interviews, which took place throughout the course of the research, focused on three primary goals: (1) ascertaining the nature of key work routines, organizational operations, program procedures, and attendant local norms; (2) developing an interpretive account of actors' understandings and their relations to choices and behaviors of interest; and (3) forging a dialogue in which our ongoing field research informed and directed our statistical analyses, and our ongoing statistical analyses raised new questions for exploration in the field.

performance pressures should provide local actors with strong incentives to make use of this information, learn from other regions, and adopt program improvements that work.

Previous studies suggest a variety of reasons why, in practice, service-providing organizations may deviate from this script in "rationally perverse" ways. In this view, performance indicators present local organizations with ambiguous cues that get "selected, interpreted, and used by actors in different ways consistent with their institutional interests" (Moynihan 2008, 9). Performance pressures often fail to stimulate positive innovations because managers lack authority to create change, learning forums are weak, and organizational routines and cultures prove persistent (Moynihan 2006, 2008). At the same time, the "tunnel vision" created by performance numbers can lead local actors to innovate in perverse ways that subvert important programmatic goals. Performance pressures may divert attention from important-but-unmeasured values and activities (Radin 2006) and, in an effort to boost their numbers, organizations may engage in "creaming" practices that direct "services to those already close to being 'job-ready' at the expense of those with barriers to employment" (Considine 2003, 71).

Our research confirms that performance concerns are deeply rooted in the consciousnesses of local WT personnel. Regional officials and program managers expect to be held accountable for the outcomes they produce; they carefully scrutinize performance reports and keep a close eye on other regions; and they express a strong desire to improve performance through evidence-based revisions of practice. Local officials routinely describe themselves as following a "business model" in which welfare-to-work outcomes are the "products" and performance is exchanged for payments.

In this regard, state officials repeatedly emphasize the need for regions to "make their bogey" (i.e., meet performance goals). Regional officials match this level of concern and affirm that performance measures drive their decision making. As one local manager explained, "We're at the bottom of the chain, and we look up to [the state level to] see what's important. And the performance measures are how we know. When you tell me I need to do participation rate, I know what my priorities are. And that's where we spend our time." Another local official summarized: "This whole process with the WT program, it's a number thing. It's about the numbers. Your participation rate, your employment rate . . .."

Interviews at all levels of the WT program suggest that performance anxieties guide thought and behavior. Yet they also point to deep contradictions in the NPM that subvert the promises of organizational learning and program innovation. Consider first the double-edged nature of competition and its relation to trust. In theory, performance competition is supposed to encourage regional managers to learn from one another's experiments. Yet this same system, when designed around comparative evaluation, also encourages managers to view other regions as competitors who have a stake in outperforming them. Our site visits make clear that the diffusion of innovations requires a modicum of trust, and this trust can be undermined by highly competitive performance systems.

Echoing others we spoke with, one local manager stated unequivocally that regions try to maintain a competitive edge by guarding their best innovations as "trade secrets" and, in this same interview, asked us not to tell other regions about new techniques being tried at her one stop. Another explained that learning from high-performing regions is also complicated by the suspicion that high-stakes performance evaluations lead "other regions" to cheat: "They can't tell you their 'best practices' because their practice is cheating [to win the] competitive game." As these statements suggest, competition works at cross-purposes with policy learning by encouraging local actors to distrust the performance numbers that other regions produce, the best practices that other regions recommend, and the wisdom of sharing their own positive innovations.

Policy learning also founders on a second dynamic in the WT program that is rooted in discursive tensions between devolution and performance management. In principle, these two aspects of the NPM are supposed to work hand in hand to promote the diffusion of best practices. Yet there is a rhetorical tension between the two. Performance reports and efforts to publicize best practices function as parts of a broader discourse, suggesting that "what worked there can work here too." By contrast, the discourse that justifies state-to-local devolution prizes locally tailored solutions and trumpets the idea that localities have radically different needs, goals, populations, and capacities. Not surprisingly, these two ideas clash in the consciousness of the local manager. When presented with performance-boosting innovations from other regions, local officials cite a litany of characteristics that distinguish the region of origin from their own. Resource differences are also sometimes cited, as in the case of one official who stated flatly: "There are best practices that there is no way we can implement or staff." The deeper tension, however, is between a discourse that denigrates "one size fits all" ideas and celebrates local uniqueness, on one side, and a system that treats localities as comparable and seeks to generalize innovations across them, on the other. One local official spoke for many when he deployed the discourse of local differences to challenge the wisdom of having regions compete on the basis of performance: "Philosophically, to me it makes more sense to compare us against us. Don't compare us against Miami. Don't compare us against Orlando. Compare us against us."

In addition to these dynamics, three additional perversities flow from tensions within the NPM between decentralized management and centralized performance evaluation. Under devolution, local actors are supposed to exercise discretion in responding to performance pressures—to draw on local knowledge to select the most effective and efficient program strategies. But program effectiveness and efficiency are not the only ways that strategies differ. From an organizational perspective, some paths of innovation are easier to pursue than others.

One such "easy path," well known to students of performance management, is to engage in creative counting of activities (e.g., Radin 2006). In the WT program, serious reforms designed to deal with problems of poverty and work are (not surprisingly) often viewed as difficult to achieve, and their performance effects are usually seen as distant and uncertain. It is far easier to change how one classifies existing activities and counts measured behaviors. As a result, as one local official told us forthrightly, "people game the numbers all the time." In describing efforts to meet the required participation rate, another regional official explained: "You have to do all sorts of things to fill the participation hours. We've got a client who we found out was taking her pastor to church on Sunday. We went out and asked her pastor to sign on saying this was community service. The trick is to find out what people are already doing and find a way to count it as work or community service. This is how you have to do it."

The search for easier organizational paths also underlies a related dynamic that we have not seen noted and will call "stream creaming." Previous studies of creaming have suggested that when performance standards are high, organizations may cope by ignoring the "hard the serve" and focusing on clients who are easier to lift above a measurement threshold (Bell and Orr 2002; Considine 2003; Dias and Maynard-Moody 2006). In an era of service integration and multiprogram one-stop centers, however, providers often hold contracts related to several programs at once. Under this arrangement, which has emerged most strongly—though not exclusively—in states where TANF programs are integrated into workforce systems, profitability in each stream depends on meeting contract "pay points" tied to performance. (As one manager put it, "If we make it [the performance standard] we get paid; then if we don't, we get zero.") Thus, the profitability of the operation as a whole can be increased by raising efforts across the board, by creaming clients in each program, *or* by strategically focusing on programs that offer "softer" pay-point targets. The latter is what we mean by stream creaming.

When state officials raise benchmarks in a program, they hope that organizations will reach for higher levels of performance in that program. But when providers are contracted for multiple program streams, they may have incentives to do the opposite—that is, to reduce their efforts in the program that now has tougher benchmarks and shift their energies to programs where pay points and profits seem easier to obtain. In discussing higher performance benchmarks recently put in place for the WT program, one local manager told us, "If they are going to make our profit closely tied to something that is so hard to fully obtain, there will be problems." Asked to elaborate, he explained that the provider would abandon efforts to meet unrealistic goals and use its resources to meet pay points in another program. In addition, the company would seek new program streams with more profitable targets. "[My employer] is committed to this welfare industry but they need to make a profit as well. We have been trying to diversify our services and clientele. For instance, we are now working with ex-offenders and things like that. We provide the support system for ex-offenders who are coming out of jail."

Finally, when local actors respond to performance pressures, they also confront "easy versus hard" paths when deciding whether to focus on improving serve to the existing client pool or, alternatively, selecting a client pool that will make it easier to meet performance goals. Evidence from all regions in this study suggests that the latter path is usually seen as easier. Accordingly, creative efforts to innovate are often directed toward reshaping the clientele rather than serving them more effectively.

In this regard, it is instructive to consider the organizational innovations that occurred in what we will call "Region A" in 2005. Early that year, representatives of Region A attended a meeting with other regional officials where they were publicly criticized for having low performance numbers. In response, regional officials decided to overhaul key features of the local operation. Acting on the assumption that low performance numbers were a result of having too many clients who were "not serious enough," Region A officials chose a path of action designed to trim the caseload down to an easier-to-serve core of clients. The relevant changes went into effect around the time of June 2005 and included the following:

1. Intake and orientation procedures were revamped so that applicants would need to attend daily classes for at least one week *before* having their application for benefits submitted. Forty hours of class attendance was required, and applicants who missed a class or showed up inappropriately dressed were required to start over the following week.

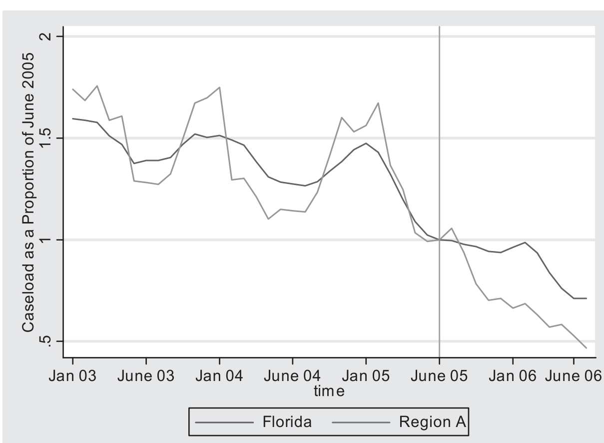
2. Intake meetings with new applicants were redesigned to emphasize the significant time investments demanded by program requirements, the limited amount of assistance available for meeting these requirements, and the fact that these requirements could be avoided if

applicants chose to pursue only Medicaid and Food Stamp benefits. As one case manager responsible for intake explained, "Doing the overview presentation, what I kinda tell them is that, if you're in a situation right now where . . . transportation is a hindrance for you, you may want to reconsider getting your cash assistance open because you're going to be required to participate in this program on a daily basis."

3. The region instituted a more frequent and intensive Quality Assurance system for monitoring caseworkers' handling of sanctions and work participation.
4. The region moved to a new system for "curing" sanctions. In the past, a sanctioned client could re-enter the caseload and reinstate benefits simply by contacting her caseworker and beginning to document work hours again. Under the new system, only one local staff member (known among the staff as the "Sanction Queen") was given authority to sign off on the return of a sanctioned client, and this staff member was made available to clients on only one day each week, for two hours. Sanctioned clients who missed this window would have to wait another week to return.

Staff from Region A consistently reported that these program innovations had major effects on the region's caseload and operations. Figure 1, based on an analysis of administrative data, corroborates this perception. Between January 2003 and June 2005, the caseload of Region A tracked closely with caseloads in other regions of the state. Both trends rise and fall in a seasonal pattern, with Region A showing slightly higher peaks and troughs than the state average. In the months immediately prior to June 2005, the two trend lines lie right one top of each other. Immediately after this date, however, the caseload in Region A falls precipitously. Indeed, despite the fact that caseloads were falling across the state during this period, Region A's caseload falls so quickly that it produces a substantial gap in relation to the state average. Moreover, although Region A's caseload had previously been

Figure 1
Caseload Effects of Organizational Change in "Region A"



more responsive to seasonal ebbs and flows, it now becomes less responsive than caseloads in other regions of the state. From June 2005 to July 2006, Region A's caseload fell by an astonishing 53%.

As a disciplinary regime, then, the performance ethos is powerful yet incomplete. It is powerful in the sense that it shapes the thinking of local officials, focuses organizational behavior, and motivates efforts to innovate at the frontlines. It is incomplete, however, because local organizations retain substantial discretion in the ways they respond to performance information and pressure. In the WT program, performance is the name of the game for local service providers. But organizations typically adapt in perverse ways, and internal contradictions embedded in the NPM work systematically against policy learning and program improvement.

PERFORMANCE PRESSURE AND SANCTIONING: A STATISTICAL ANALYSIS OF ADMINISTRATIVE DATA

The operational shifts observed in Region A suggest that performance pressures can influence local implementation in ways that severely limit access to assistance for low-income families. Building on this observation, we turn now to the question of how performance pressures affect the use of sanction procedures that limit access to assistance for explicitly disciplinary purposes. Do local providers respond to performance pressures by imposing sanctions on WT clients at higher rates? Prima facie evidence of a relationship is suggested by two facts noted earlier: Florida has one of the strongest performance systems in the country and it sanctions clients at a higher rate than any state studied by researchers to date. Because we lack consistent indicators of performance pressures and sanction rates across the states, however, we cannot conduct a state-level test of whether this correspondence is more than circumstantial.

Instead, our hypotheses in this section specify a series of "observable implications" that should be discernible in administrative data from the Florida WT program if performance pressures drive sanction rates upward. Because our field research suggests that performance anxieties are pervasive in the WT program, we begin with two hypotheses related to "chronic" performance pressures—that is, pressures felt throughout the regions on a more or less continual basis. The first takes advantage of program change over time. In July 2000, Florida shifted from the less performance-focused WAGES program to the more performance-driven WT program.[3] Accordingly, we specify:

The Program Hypothesis: Because performance monitoring, feedback, and pressures increased with the transition from the WAGES to the WT program, we expect statewide sanction rates to increase after July 2000.

3 With the onset of the WT program, Florida implemented a new system focused on tracking performance information, One-Stop System Tracking, as well as stronger procedures for monitoring, disseminating, and rewarding/penalizing performance. The sanction rates shown for the WAGES program in figure 2 cover a comparatively short time span because no comparable data are available for sanction rates prior to January 2000.

If chronic performance pressures contribute to higher sanction rates across the state, then one should also observe higher sanction rates in service-providing organizations that are more sensitive to performance incentives. Since its inception, the WT program has been implemented through contracts with a variety of nonprofit and for-profit organizations. In theory, the key difference between the two should be the ‘‘nondistribution constraint’’—that is, the inability of nonprofit organizations to distribute profits to managers and shareholders. Because of this constraint, some predict that for-profit managers will stress performance incentives and pay points to a greater degree, whereas nonprofit managers will be less concerned with cost minimization and, hence, more likely to ‘‘expend resources to serve segments of the public that would otherwise be seen as too costly or unprofitable to serve’’ (Heinrich 2000). Although other scholars point out that both types of organizations are sensitive to revenue concerns (Weisbrod 1998), the distinction remains strong enough at the level of theory to specify:

The Organizations Hypothesis: Under the strong performance pressures of the WT program, sanction rates will be higher in for-profit than in non-profit service-providing organizations.

Although the most significant feature of performance pressure in the WT program is its ubiquity, more precise tests are possible if we examine how local sanction patterns respond to episodic changes in performance feedback over time. The pervasiveness of performance anxieties in the WT program should produce a ‘‘ceiling effect’’ that makes it difficult to observe large short-term changes in sanction rates in response to the periodic publication of performance numbers. Yet, as we have seen, regional officials say they pay close attention to performance reports and are eager to respond in ways that might improve their numbers. Accordingly, we test three ‘‘episodic’’ hypotheses.[4]

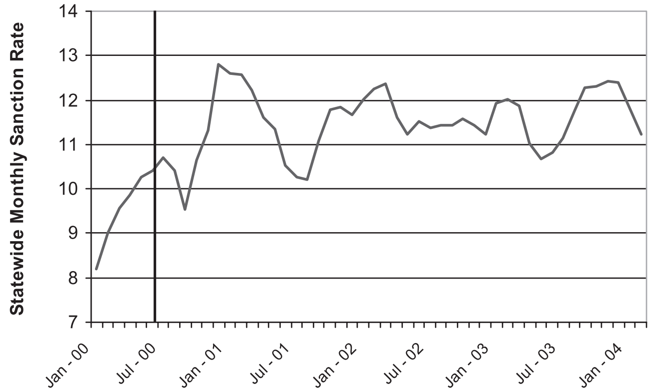
The Feedback Hypothesis: Because regional officials and case managers experience increased performance pressure in the wake of negative performance feedback, we expect regional sanction rates to increase in response to declining regional performance.

The Clientele Hypothesis: Because the WT system penalizes regions that are less successful in moving clients into the workforce, we expect negative performance feedback to raise sanction rates to a greater degree for clients who are ‘‘harder to serve’’ or viewed as lacking work motivation.

The Ideology Hypothesis: Because providers in politically conservative locales tend to rely more heavily on sanctions to motivate clients (Fording, Soss, and Schram 2007), we expect negative performance feedback to stimulate sanctioning to a greater degree in more conservative regions.

4 Because we have been unable to obtain longitudinal data on local one-stop operators, we are unable to test an episodic variant of our organizations hypothesis—that is, whether for-profits and nonprofits vary in the ways they respond to declining performance feedback over time.

Figure 2
Change in the Welfare Sanction Rate over Time: The Transition to WT



To test the Program Hypothesis, figure 2 tracks the average sanction rate for Florida workforce regions from January 2000 to April 2004, based on a 3-month moving average of the percentage of the adult caseload receiving a sanction. The trend line indicates a seasonal pattern to sanctioning, which complicates interpretation because rates were on the rise just prior to the inauguration of the WT program. On the whole, though, the trend line is consistent with our expectation. The statewide sanction rate rises significantly after July 2000 (denoted by the vertical line in the figure) and, under the WT performance system, never falls again to its lower rate under the WAGES program. The average sanction rate from January to June 2000 was 9.02%. Taking the seasonal pattern into account, we find that the average sanction rate for the same 6-month period (January–June) for years 2001–2003 was 11.8%, a percentage increase of roughly 31%.

To test the Organizations Hypothesis, we use a sample consisting of all adults who participated in the WT program for at least 1 month during the period, November 2003–April 2004. Our dependent variable indicates if a client was sanctioned during this period (1) or not (0). Our key independent variable equals 1 for the 16 regions where one-stop centers were operated by for-profit firms and 0 for the 8 regions that contracted with nonprofit organizations. We control for a number of variables suggested by the literature on sanctioning, including client characteristics and measures of local economic and political context (Fording, Soss and Schram 2007; Pavetti, Derr, and Hesketh 2003). The results presented in table 1 indicate that, even after controlling for these conventional factors, WT clients were significantly more likely to be sanctioned if they participated in regions where one-stop centers were operated by for-profit firms. Indeed, the odds of receiving a sanction are estimated to be 25% higher for such clients, relative to similar clients served by nonprofit providers.

To test our three episodic hypotheses, we employ a panel data set consisting of aggregate monthly observations for each of Florida's 24 Workforce Board regions. This data set consists of 30 monthly observations for each region, spanning the period, October

Table 1
Effect of Organizational Form on Individual Sanctioning Outcomes

Independent Variables	Odds Ratio
Marital Status (1 = unmarried)	0.88**
Black	0.98
Hispanic	0.99
Number of Children	0.97*
Age of Youngest Child	1.00
Gender (Male = 1)	1.23**
Citizen	1.17*
Age (years)	0.98**
Spell Length (months)	1.08**
Local Conservatism	1.06
Local Unemployment Rate	1.03
Per Capita Caseload	0.61
For-profit Provider	1.25*
Log Psuedo-likelihood	−32786.595
Wald c^2 (13)	749.99**
Sample Size	53,373

Note: The coefficient entries are odds ratios, based on a logit analysis. Significance tests are based on robust SEs, adjusted for clustering by county of residence.
**$p < .05$, *$p < .10$, two-tailed test.

2001—March 2004.[5] Our measure of regional performance is based on the average (monthly) regional ranking (1–24) across the three key measures used by state Workforce Board to monitor regional performance in the WT program—the entered employment rate, the welfare return rate, and the entered wage rate.[6] We sum the rankings of these measures so that higher scores indicate *stronger* pressures (i.e., declining performance).

The dependent variable for these analyses is the regional sanction rate, which we define as the percentage of each region's monthly caseload that received a sanction in a given month. The control variables include a variety of characteristics of the adult caseload, including racial/ethnic composition, average age, the work participation rate, family structure, family size, TANF dependency, and the overall size of the monthly adult caseload. (Definitions and descriptive statistics are presented in table A1.) Finally, each analysis includes a full set of regional and monthly fixed effects. The former control for all time-invariant factors that vary across regions, thus providing control for unmeasured

5 This time period reflects the maximum amount of time for which we are able to obtain data for regional performance rankings, regional sanction rates, and the characteristics of TANF clients.

6 We use the regional ranking, rather than the actual performance measures themselves, because performance incentives are largely based on the region's performance relative to other regions in the state. We measure average regional ranking cumulatively within each fiscal year (i.e., the region's average performance ranking from the first month of the fiscal year [July] through the most recent month) because performance incentives are based on a region's performance across the entire fiscal year. Specifically, we measure regional performance as the monthly *change* in the average ranking. This reflects our belief that local TANF administrators are likely to be most responsive to the short-term trajectory of the region's performance ranking, rather than the overall ranking itself. Finally, our measure of performance feedback is lagged 2–3 months to account for the lapse of time between the end of the month, the publication of the monthly performance reports by the state (which are not available until the following month), and the communication of the regional response by regional managers to frontline personnel.

Table 2
Effects of Regional Performance Measures on Regional Sanction Rates

Independent Variables	All Clients	Race of Clients: Black	Race of Clients: White	Clients' Education: <12 years	Clients' Education: ≥12 years	Political Ideology: Most Liberal	Political Ideology: Conservative
Performance Ranking	0.127**	0.200**	0.137*	0.234**	0.173**	0.017	0.180**
Participation Rate	−0.074**	−0.126**	−0.033	−0.101**	−0.093**	−0.124**	−0.037
Average Time on Current Spell	−0.873**	−0.357	−1.205**	−1.807**	−0.642	−1.967**	−0.437
Average Spell Number	−1.877	−0.962	−2.707*	−2.227	−1.798	−0.965	−5.552**
Average Number of Children	3.980*	1.738	2.392	3.066	.909	.189	3.910
Average Age of Youngest Child	−0.181	−0.047	0.808	−0.375	0.096	−1.746**	1.274
% Single	−0.069	−0.054	−0.046	−0.053	0.077	0.049	−0.043
% Male	0.061	−0.035	0.119	−0.072	0.098	0.012	0.198
% Citizens	−0.180	−0.339*	0.077	0.024	−0.050	0.142	−0.464**
% Black	−0.147**	—	—	−0.036	−0.122**	−0.238**	−0.140
% Hispanic	−0.166*	—	—	−0.005	−0.125	0.020	−0.333**
Average Age of Adult Clients	−0.316	−0.878**	−0.191	−0.031	−0.561	0.311	−0.792
Size of Caseload$_{t-1}$	−0.450	−0.661	−1.328*	−2.161**	−0.282	−1.117*	1.249
Sanction Rate$_{t-1}$	0.271**	0.240**	0.137**	0.131**	0.084	0.241**	0.245*
R^2	0.58	0.48	0.46	0.52	0.49	0.68	0.48
Sample Size	N = 720	N = 720	N = 720	N = 552	N = 552	N = 360	N = 360

Note: The dependent variable is the monthly sanction rate (no. sanctioned/caseload × 100). *Performance Ranking* is defined as the change in the average monthly ranking (1–24) for the entered wage rate, the welfare return rate, and the entered employment rate. This variable is measured at a lag of 3 months and is measured cumulatively within each fiscal year. Slope coefficients are estimated by ordinary least squares, whereas significance tests are based on panel corrected SEs. All models are estimated with a full set of fixed effects for workforce regions and month of analysis.
**$p < .05$, *$p < .10$, two-tailed test.

differences in regional contexts that may affect client outcomes, whereas the latter control for time-varying variables that do not differ across regions, such as changes in state-level policies that affect all regions.

We begin with a test of the Feedback Hypothesis. The results, presented in the first column of table 2, are consistent with our expectations: for each one-unit increase in our ranking-based measure of performance pressure, the regional sanction rate increases by an average of about 0.13.[7]

The Clientele Hypothesis suggests that negative performance feedback should increase sanction rates to a greater degree for harder-to-serve clients. Because black clients, relative to white clients, confront more barriers to labor market success (Holzer and Stoll 2002) and are subject to stronger stereotypes of preferring welfare to work (Schram et al. 2009), this hypothesis suggests that negative performance feedback should increase

7 Because we include a lagged dependent variable in our models, this estimate represents only the immediate effect of performance feedback, with the effect distributed over time through the lagged dependent variable. Even so, this effect is statistically discernible but modest. After 6 months, the cumulative effect stands at 0.17.

sanction rates for black clients more than white clients.[8] Similarly, we should expect negative performance feedback to increase sanctions for clients with less than 12 years of education to a greater degree than for clients with 12 or more years of education.[9] The results of these analyses, presented in columns 2–5 of table 2, generally support our expectations. Performance feedback has a statistically significant effect in all client categories, but the effects are significantly larger among black clients and less-educated clients. The effect among the less educated exceeds the effect among the more educated by approximately 35%. The effect for black clients is 50% larger than for white clients.

The Ideology Hypothesis suggests that providers in conservative regions, that tend to embrace a stricter welfare-to-work approach, are more likely to turn to sanctions as a response to poor performance feedback. Local environments may affect organizational operations through democratic pressures because officials respond to local conditions and needs or because officials share the political values of the community (Goggin et al. 1990; Weissert 2000).[10] To measure local political variation, we use an index of regional conservatism described in table A2. The results, presented in columns 6–7 of table 2, indicate that performance feedback has no discernable effect on sanction rates in the 12 most liberal workforce regions. In the 12 most conservative regions, the effect is significant and roughly 10 times what we observe in the most liberal regions. In conservative regions, a one-unit increase in our ranking-based measure of performance pressure leads to an immediate increase in the sanction rate of approximately 0.18. After 6 months, the cumulative effect is modestly larger, estimated to be 0.24.

Extending the logic of our separate hypotheses, one might also expect an interactive effect in which more conservative regions disproportionately sanction hard-to-serve clients in response to negative performance feedback. In separate analyses not shown here (available on request), we find results that are largely consistent with this expectation. In liberal regions, we find that the relationship between performance feedback and sanction rates does not vary significantly across client subgroups. In conservative regions, however, we find that sanction rates increase in the wake of negative performance feedback in a pattern that disproportionately targets harder-to-serve clients. Specifically, they increase to a significantly greater degree for black clients and clients with low education levels. These results suggest that the Clientele Hypothesis is more accurate for conservative regions than for liberal ones.

Finally, as our earlier discussion of "Region A" illustrated, service-providing organizations can employ a variety of strategies beyond sanctions to limit and shape their caseloads. Indeed, regional officials may respond to negative performance feedback by pursuing restrictive measures that are not reflected in sanction rates. The case closure rate—calculated as the percentage of open cases each month that are closed for reasons other than sanctions or earnings—provides a way to capture the sum total of such efforts to remove cases in the face of performance pressure.[11] If local TANF offices do indeed

8 Because the Hispanic caseload is concentrated in a limited number of regions across the state, we cannot perform an equivalent panel analysis and, hence, compare performance-feedback effects for black and white clients only.

9 Limited data on client education levels force us to restrict the time period for this analysis to 23 months.

10 We also tested for an interaction between our measure of local ideology and for-profit status, as an elaboration of the model reported in table 1. The coefficient for the interaction term was in the predicted direction (positive); however, the significance level fell short of conventional standards (0.14).

11 In the analysis that follows, our dependent variable is the natural log of the case closure rate, as published in the state of Florida's official regional caseload reports.

Table 3
Effect of Regional Performance Measures on Regional Case Closure Rates

Independent Variables	All Regions	Liberal Regions	Conservative Regions
Performance Ranking	0.007*	0.003	0.012**
Participation Rate	0.001	0.004**	−0.003*
Average Time on Current Spell	0.005	0.093**	−0.012
Average Spell Number	0.119	0.464**	0.299*
Average Number of Children	−0.023	0.004	0.092
Average Age of Youngest Child	0.056	−0.013	0.117*
Average Age of Adult Clients	0.027	0.013	−0.000
% Single	−0.001	−0.011	−0.001
% Male	−0.011	−0.028**	−0.004
% Citizens	−0.008	−0.012	−0.022*
% Black	−0.005	−0.007	−0.001
% Hispanic	0.004	0.006	−0.001
Size of Caseload$_{t-1}$	0.013	0.084**	0.021
Case Closure Rate$_{t-1}$	0.099*	0.162**	−0.047
R^2	0.51	0.68	0.39
Sample Size	$N = 720$	$N = 360$	$N = 360$

Note: The dependent variable is the natural log of the monthly case closure rate (no. cases closed/caseload × 100). Cases closed in this measure are for reasons *other* than sanction or earnings. *Performance Ranking* is defined as the change in the average monthly ranking (–24) for the entered wage rate, the welfare return rate, and the entered employment rate. This variable is measured at a lag of 2 months and is measured cumulatively within each fiscal year. Slope coefficients are estimated by ordinary least squares, whereas values in parentheses are panel-corrected SEs. The period of analysis is October 2001–March 2004 (30 months). All models are estimated with a full set of fixed effects for workforce regions and month of analysis.
**$p < .05$, *$p < .10$, two-tailed test.

respond to declining performance by shedding clients, this response should be manifested, not only in the sanction rate but also in the rate at which clients are pushed off for other administrative reasons.

The results, presented in table 3, corroborate our analysis of regional sanction rates. The effect of performance feedback is statistically significant for the full sample of workforce regions (column 1), and the slope coefficient suggests that for each increase of one unit our measure of performance pressure, a region's case closure rate is expected to increase by 0.7%. Due to data limitations, we are unable to test the Clientele Hypothesis with case closure data.[12] For the Ideology Hypothesis, however, we find a pattern similar to the one reported for sanction rates. Performance feedback has no discernible effect on case closures in the most liberal regions, but in the most conservative regions, the effect is four times as large and statistically significant. For each unit increase in our performance measure, the case closure rate increases by 0.012 in these regions. These effects are smaller than for the sanction rate, but the consistency of findings across two dependent variables should bolster our confidence in these results.

12 For non-sanction exits, the individual-level data we use to build our control variables for racial and educational subgroups do not distinguish between clients who leave TANF due to earnings and clients who exit for other reasons. However, such data are available to us at the regional level for the entire caseload.

Our quantitative results consistently match what we would expect to find if performance pressures motivate service providers to push clients off the rolls more frequently. Under the strong performance system implemented in 2000, chronic performance pressures have raised sanction rates across the state. Sanction rates under this system are slightly but significantly higher among the for-profit providers we would expect to be especially attuned to the profit implications of performance-based pay points. Episodic declines in performance rankings also appear to heighten performance pressures in ways that lead to increases in sanctions and case closures. Moreover, these effects vary across client subgroups and regions in the ways we would expect if performance pressures have a real impact on sanctioning. Although these effects are often modest, their consistency gives us greater confidence in this conclusion.

FROM PERFORMANCE PRESSURES TO SANCTIONING: A FIELD PERSPECTIVE ON MECHANISMS

With this evidence in hand, we turn now to the question of *how* performance pressures influence case managers' sanction decisions. The most logical candidate for a causal mechanism in this context, and the one most clearly suggested by the literature on perversity, is the practice of "creaming" (Bell and Orr 2002). In response to performance pressures, frontline workers might use sanctions as a strategy to rid themselves of low-performing clients and thereby restrict their caseloads to the "cream" clients who generate positive numbers. In addition to being a well-established hypothesis, this expectation fits well with the "perversely rational" behaviors we found in our analysis of organizational responses. Thus, when we started our research on case manager discretion and sanctioning, we expected to find a creaming dynamic. Our field research, however, failed to cooperate with this expectation. Today, we refer to the strategic creaming account as the "causal story that failed."

Why? To begin with, our interviews made it clear that case managers are rarely single-minded performance maximizers. More typically, they are ambivalent actors caught in the cross-pressures of competing values, identities, and organizational forces (see also, Watkins-Hayes 2009, 2011). Despite the rhetoric of the "business model," most express a strong commitment to social service ideals and value their identities as providers who are responsive to clients' needs. As a result, they express deep reservations about making case decisions based on performance goals. Consider the following quotations.

> They say that we're not, how would you say it . . . a social service agency in a sense, like we're business . . . But at the same time . . . you're working with people who have needs, who have barriers, and bringing the two together is very difficult. [. . .] There's a number game that we have to play. And when you bring that into it, it's hard for me to sit with an individual there; they're telling me that they have all these barriers. For example, they're coming in and they're telling me that they've been evicted from their apartment, they don't have any place to live, they don't have any food, they don't have any clothes. And then here I am as a case manager, you have to participate at 40 hours a week. You know, it's just kind of, its crazy!"
>
> The way we're able to [stay in business and] help people is by making our measurements on our red and green reports and getting paid, so that we can therefore in return help with childcare and support services [. . .] So the more we make those measurements and those goals, the more we can help candidates. But the more we focus on those [performance goals], the less we're focusing on the candidates. So, it's a catch-22.

Ambivalence regarding performance management is matched by ambivalence about sanctions themselves. Like performance indicators, sanctions are central to the ways that case managers in the WT program perceive and pursue their mission. As one official put it, ''sanctions are the most important process we have in terms of case management and in terms of producing results.'' Senior officials and case managers are also virtually unanimous in supporting sanctions, in principle, because they believe there should be consequences for client noncompliance. As one case manager put it, ''I think realistically, you have to have teeth in the program to get people to participate.'' Yet at the same time, many have doubts about sanctions, in practice, because they worry that clients are set up to fail: they are required to meet tough program requirements without the supports and services they often need to do so. Thus, many say that sanctions are ''educational tools'' and ''not punitive at all,'' yet this view lies uneasily beside the belief that ''Florida has a punitive system that gets increasingly harsh the more problems a [client] has. [A sanction] is not a deterrent at all. It's used like a punishment.'' Reflecting this ambivalence, one official confidently asserted that sanctions are the best way to ''stop [clients] from wasting their own time'' but later hedged, ''To be honest, I'm not always sure what sanctions are good for. Sanctions are just a reality of the program. They don't really deter or gain more attention. They're just how the program works.''

Such widespread ambivalence makes it difficult to maintain the view that case managers intentionally sanction clients to advance their narrow-minded pursuit of good performance numbers. The creaming account, however, is contradicted more fundamentally by a second observation. Throughout our interviews, we found strong and nearly universal adherence to one basic belief: in the WT program, high sanction rates are *bad* for performance. Thus, even if case managers were single-minded performance maximizers, they would be unlikely to pursue a creaming strategy because *virtually no one in the Florida WT program believes that sanctions have a positive effect on performance rankings*. Regional officials consistently state that high sanction rates hurt performance and invite unwanted attention. As one put it, ''Our region doesn't want to have a sanction rate that's too high. High numbers (of any kind) draw attention to the region, so it had better be something positive. So we wouldn't want to be seen as overly punitive in a way that might not be within the rules.'' Case managers consistently told us that ''if sanctions get high they hinder [our numbers],'' often adding that ''[supervisors] want you to maintain your sanctions as low as possible.'' Significantly, this message from supervisors is couched, not just in terms of performance, but in terms of the values and goals of the business model. As one explained: ''This is a private company, and our goal is to get them employed, not sanctioned.''[13]

Instead of a ''creaming'' dynamic, our findings point to a more subtle explanation of how performance pressures influence sanctioning. Our account, built inductively from field interviews, underscores the importance of an organizational perspective that highlights how case managers are constrained and disciplined by the environments in which they are embedded. The most plausible mechanisms can be traced, we believe, to the

13 Confidence that sanctions hurt performance should not be confused with clarity about how, precisely, the two relate. At numerous meetings, we observed regional staff disagreeing about how sanctions factor into performance calculations and asking state-level officials for clarification. When asked for *details* about how sanctions affect specific performance measures, case managers frequently laughed and said that they honestly were not quite sure.

conjunction of four factors: (1) the distinctive ways that the WT program organizes case management, (2) the specific performance pressures experienced by case managers, (3) the limited number of tools available to WT case managers, and (4) case managers' beliefs and frustrations regarding client noncompliance.

The path from performance pressures to sanctions ultimately runs through the organization of WT casework, which is highly routinized and focused on performance-related tasks. Caseworkers typically describe themselves as responding to system needs rather than acting proactively. The following exchange among senior officials underscores this point.

> Regional Official A: You don't hire a "people person" anymore for a career manager position. You hire a clerical computer person. You can teach them the social work stuff easily. The job's all about time, accuracy, and files now. There's a person [client] down there somewhere. But the technical stuff is what matters.
>
> State Official: What you're telling me is the [information] systems are driving the [case management] process.
>
> Several Regional Officials: Oh yes. Oh yes!
>
> Regional Official B: You don't get any credit [in the performance measures] for hand-holding. You don't get any credit for mentoring. [. . .]
>
> Regional Official C: If you talk to any case manager here, they will tell you they're not a case manager; they're a technician. They spend about 10 percent of their time on their clients. Their time is about being a technician, and that's the way the program is written. They're doing what they have to do under this system.

WT case management is reactive and clerical. It focuses primarily on documenting client activity hours and entering the results into the One Stop Service Tracking (OSST) data system. Indeed, managers at several levels argued that the data-entry fields of OSST function, in daily organizational routines, as the real policy on the ground. As one put it, "The policy [on the books] doesn't always match up with the [OSST] system. People on the frontlines see the computer screens as the policy. Whatever can or can't be done in a straightforward way on the system, it's assumed that that's the policy."

When asked to describe their workday, case managers consistently report that they begin by logging on to the information system so they can address the slew of new alerts that arrives each morning. The alerts focus on two kinds of actions: documenting work participation hours for clients and pursuing disciplinary actions when such documentation is lacking. From this point forward, the daily round consists mostly of efforts to do one or the other, punctuated by face-to-face meetings with clients that often focus on the same two issues. Case managers spend most of the day either seeking documentation for work-related activities (a key performance indicator) or taking next steps in the sanction process such as sending out a "prepenalty" warning letter, requesting a sanction, or working to bring a sanctioned client back into compliance. In short, performance and sanctioning are two sides of a single coin in the work life of the case manager and, together, they stand at the center of the job.

As a result, case managers worry about performance almost continually. As one put it, "It's just weird, I mean it really is. And I don't know how to explain how, um, you know, we [case managers] all run around and we're like, 'where are you at now with your [participation numbers]?' 'Oh, I'm at like 20 percent.' 'Oh man!' So we're all just stressed!"

The stress felt by case managers can be traced partly to their belief that performance numbers matter for job security and trajectory. WT case managers make modest wages in a job with few guarantees, and a nontrivial number have previously received welfare themselves. They often struggle to make ends meet and, as a result, tend to view performance through the prism of their own anxieties as breadwinners. Few expect to be "fired" if their numbers drop. But in a system of for-profit contracting, most are keenly aware that performance numbers drive profits, and declining profits could lead their current employer to downsize the staff or even to sell the operation to another company whose retention of old employees is uncertain. At a less absolute level, most expect that if they produce weak numbers, they will be subjected to greater supervision in a way that will make their work more stressful and harder to do. One case manager explained, "We [case managers] get our own sanctions. [Laughs] So, um, you know, that's a big stress. Um, and they also tell us, 'yeah, the entered employment; um, how many jobs are you getting?' [. . .] I mean, that's just things that are hit every day, fifty percent [participation], fifty percent."

In describing these pressures, case managers make it clear that they do not see sanctioning as a desirable response. In addition to expecting negative effects on performance numbers, most caseworkers do not like to impose sanctions, and many are skeptical that sanctions have positive effects on clients. One explained, "No career manager wants to sanction. You go through all these papers to try to get in touch with the person. It's a struggle. You try to help them get everything in to keep them out of a sanction, but a lot of times you can't."

Despite this resistance, our field research makes it clear that sanctions are central to the ways that case managers respond to performance pressures because, to put it simply, they have few alternative tools at their disposal. They are limited in what they can do to raise their numbers, and they are even more limited in their abilities to address the real-life problems of their clients. Most caseworkers have no training as social workers, they have few options for matching clients to services, they are essentially powerless to change clients' opportunities and life conditions, and program rules make sanctions into a default response when cases do not fit easily within the rules and tools of the program. Buffeted by performance pressures and lacking the tools to respond to client needs, case managers experienced their workdays as a series of frustrations and disappointments. The following quotations illustrate:

> Never mind that Deborah can't read, and she's got a 6th grade education, but you want [her to] go out and get a job at ten bucks an hour. Or, my candidate, who has a substance abuse problem, you know, he keeps drinking on the job, that's why he can't *keep* his job, but [he's] got to go out there and get a job, you know [. . .So] I think we're more frustrated about meeting our participation rate every single month—that [fifty] percent. [. . .] It's a big frustration because you're like "I want to make my fifty percent. I don't want to be evaluated at the end of the month and told, "Oh, you didn't make fifty percent." What do you do?
>
> We try the best not to sanction clients, try to help them overcome barriers. [. . .] But if we follow regulations and procedures – if they [clients] don't do their part – we have to go by what the program says [and impose that sanction]. [. . .] The program regulations need to be looked at immediately [at the] highest level and [they need to] give us the tools to help clients to a better outcome.

Frustrated and lacking effective policy tools, case managers must find ways to reconcile competing demands. A small number square this circle by shifting part of the client's burden onto themselves. They put in long hours to establish whatever documented activities are needed to satisfy performance goals and avoid a sanction. With little time to spare at the office, they often spend hours at the end of a long workday trying to locate clients and secure documentation. As one explained, "My level of sanction is so low. If I were to go by policy [alone], all of [my] caseload would be sanctioned. [. . .] I go way and beyond [the policy]: a lot of communication, a lot of calling, trying to find where they're at. 'Hey, this is what's going on. If you don't come in, your benefits are going to be stopped." Going further, some case managers in this group make use of "creative counting" strategies as a way to soften the program rules confronted by clients. Here again, we see how the disciplinary power of performance management occasions its own resistance (Foucault 1980). But it is a subversion that comes at a substantial personal cost. Case managers in this group report that they are exhausted, burned out, and disappointed that their job is so often about protecting clients from the program itself.

For most WT case managers, performance pressures are a more controlling organizational reality, shaping their use of discretion. Ultimately, most believe that it is the client alone who is responsible for documenting work activities, and it is the client who must, in some way, be prodded to do so. The problem is that, aside from the threat of sanctions, case managers possess few tools for motivating clients. In principle, incentives for good behavior, such as childcare vouchers or transportation assistance, offer an alternative disciplinary tool for shaping clients' calculations regarding desirable behavior. In practice, however, appeals to these incentives often amount to little more than an implicit threat of sanction. Most benefits for program participants are already available to the client. So the discussion is, at root, about the possibility of a current benefit—or a future transitional benefit—being terminated through a sanction. Thus, as one official explained, sanctions are usually the most effective "tool for helping clients see the benefits of sticking with the program in order to get transitional benefits."

Although most case managers would like to avoid imposing a sanction, they quickly find themselves turning to the *threat* of a sanction as a way to cajole clients into participating at higher rates and turning in their documentation. When such threats fail, as they often do, case managers find themselves initiating "pre-penalty" actions as a way to signal that they "mean business" and will impose a sanction if the client does not do what is required. At this point, organized sanction procedures are set in motion. The computer alerts and requests for action kick in and the caseworker's discretion over the case diminishes. If the client now fails to comply, the case manager confronts strong pressures to move the sanction process forward in a "timely manner." As one explained:

> If you have a customer who turns in fifteen [hours] instead of the 20, at that point in time you can elect to start the penalty process or you can elect to call them and say "Okay, you know what, you have 20 due. It could have just been an oversight or whatever. Can you bring in the other five today, by 5:00 p.m.?" So you have a little flexibility to work it. *But once you start that [pre-]penalty, that 8-10 process, there's no way to work around that.* It's going to pretty much take its course. Okay, you place the [pre-]penalty on the customer's case; you give them the pre-penalty call. Try and find out, okay, why have you failed to participate? Um, they have the 10 days to comply. [If there's] failure to comply, it will lead to a sanction.

This, then, is the first mechanism that explains how performance pressures increase sanction rates. It is a story of intentional tactics producing unintended outcomes. With few tools at their disposal, caseworkers turn to sanction threats in the hope that compliance will ensue, performance numbers will improve, and a sanction will be averted. Once set on this path, though, they must (however reluctantly) put one foot in front of the other. In a short time, they find themselves imposing the sanction that, in the abstract, they see as a hindrance to high performance.

There is also a second mechanism at work, and it is a dynamic that one might expect few case managers to reveal. In a surprising number of instances, case managers report that they sanction clients out of frustration. Performance pressures contribute to this dynamic in two ways. First, as noted above, performance pressures combine with limited tools to produce high levels of frustration at the frontlines. Second, the performance system is structured so that evaluations of the case manager depend on client behavior. When clients fail to turn in documents on time, for example, their actions lower the case manager's performance numbers and invite unwanted scrutiny from supervisors. Not surprisingly, clients tend to become the focal point for case managers' frustrations in such instances. (And it is worth recalling here the extent to which case managers believe that their jobs depend on performance numbers.)

In the WT program, the noncompliant client is not just behaving in a way that *concerns* the case manager; she is doing something *to* the case manager. As one case manager explained, "The stress is, okay, well *I'm* caring about this, but the customer doesn't care. So then after a while, you still do what you got to do because you need your job, and you got to make your [measured] hours." Another reported:

> When it comes to, you know, the problem cases, we get frustrated. I think some people say, "Yeah, technically I could give her another day [to get her documents in], but you know what, I'm gonna slam it [a sanction] on her." You know? [Laughs . . .] It's that whole accountability thing. Because *we* have to be accountable, so I think when you get a customer that doesn't feel that they have to be *as* accountable, you can get frustrated.

This frustration-sanction dynamic was prevalent enough to be openly discussed at statewide training sessions. In explaining the role of sanctions, one trainer began by observing that "some people want to penalize because they're angry with a client. That's not the point." In a private interview afterward, a state-level official elaborated: "There [is] no training about case management or emotional issues. Anger management is a big issue. Case managers snap and then sanction because they're mad."

In sum, then, the causal processes that link performance pressures to sanctioning do not run through the strategic rationality of individual case managers seeking to maximize their numbers. They are more deeply embedded in the organization of case management itself. As performance pressures rain down on the frontlines, case managers are positioned as their ultimate repository yet are given few effective tools for responding. Lacking alternatives, case managers turn to the most basic threat they can wield—the sanction—as a way to motivate client compliance. Predictably, the threat leads to a first procedural step, and what was at first intended as saber-rattling turns into a sanction. At the same time, many case managers become frustrated with clients and perceive an injustice in the fact they are being held accountable while the client is not. In such circumstances, it is not hard to see why case managers often "snap" and levee a sanction that, under other circumstances, they would prefer not to impose.

CONCLUSION

In poverty governance today, performance systems and the NPM are shrouded by free-market images of autonomy, innovation, and efficiency. They are rarely seen or investigated as disciplinary regimes. Yet performance systems discipline frontline officials just as surely as sanctions discipline clients. Indeed, the technologies of discipline that govern actors on the two sides of the welfare case manager's desk have strong parallels. Both rest on incentives for right behavior and penalties for noncompliance; both aim to reshape the motivations of targets so that they will pursue preferred ends as self-regulating subjects; and neither controls behavior completely enough to forestall subversion. Just as welfare clients resist and evade the supervisory regimes of welfare-to-work programs (Gilliom 2001), so too do service providers subvert the goals of performance management at the frontlines of welfare reform. As Foucault (1980) notes, this ubiquity of resistance should not be confused with a weakness of disciplinary power. To the contrary, performance pressures have profound effects on consciousness and behavior at the frontlines of welfare reform, and these effects matter greatly for the disciplinary penalties that are meted out to the poor.

In this sense, our analysis has important implications for how scholars understand street-level discretion. Images of uncontrolled case managers acting on personal whim have been a staple of anxieties regarding welfare reform, even from its inception. Critics on the right have worried that liberal or lazy frontline workers might not really implement the demanding new procedures of welfare-to-work programs. Critics on the left have worried that expanded program requirements and tough punitive tools would give case managers *carte blanche* to treat clients in arbitrary and unjust ways. Field studies of implementation have, at times, inadvertently reinforced these anxieties by making it clear that frontline discretion is endemic, cannot be eradicated by supervision or procedure, and functions to rewrite policy on the ground as street-level bureaucrats select, interpret, and adapt the broad rules they inherit.

Our findings do not contradict this conventional wisdom or the studies that have advocated it (e.g., Brodkin 1997; Lipsky 1980; Maynard-Moody and Musheno 2003). Rather, they underscore the perils of taking this lesson from the literature while failing to appreciate what leading scholars have equally emphasized: the organizational forces that shape worker autonomy and channel behavior at the frontlines. As scholars such as Lipsky (1980) and Brodkin (2007, 2011) have repeatedly argued, the fact that frontline workers are weakly constrained by rules does not mean that they are free to act as they wish. Their uses of discretion are not "*ad hoc*, unsystematic, or incomprehensible" (Feldman 1992, 163), nor are they mere reflections of individual preference and decision making (Baumgartner 1992, 129; Mashaw 1983, 213). Indeed, a central point of the literature on street-level bureaucracy has always been that organizational routines, tools, norms, incentives, information systems, and categories of understanding function as mechanisms of social control that shape the use of discretion in predictable ways.

Our analysis of performance management underscores the limits of discretion in the work lives of local program managers and caseworkers. In the Florida WT program, the discretion possessed by case managers is *broad*, in the sense that they are authorized to make a wide variety of decisions affecting the client, and it is *ineradicable*, in the sense that they almost always know some way to push a decision in a preferred direction. But case manager discretion does not run very *deep* if by "deep" one means an individual liberty to treat clients as one would like. WT case managers are under tremendous performance pressure and have strong incentives to attend to this pressure. The computer key strokes needed to initiate action, by design, ensure that their uses of discretion will be tracked in the information system. And partly as a result, case managers make their choices as actors who know they are being observed and evaluated. Interviews with WT case managers suggest that welfare reform has initiated a tougher regime of social control, not just for welfare clients but also for the employees who serve as their caseworkers. Indeed, our analysis suggests that high rates of sanctioning in the WT program are more closely tied to this organizational regime than to any expectations case managers have about how sanctions affect clients or performance numbers. Strong performance pressures promote the use of sanctions to discipline the poor because they are a form of coercive power that drives and directs action and because they are a form of productive power that shapes subjective understandings, perceptions, and choices at the frontlines.

To be sure, our analysis also points to a substantial amount of discretion in street-level organizations. We see it, for example, in the case managers who shift burdens onto themselves in an effort to protect clients, who act on their personal frustrations by sanctioning in cases where they otherwise would not, and who use threats and impositions of sanctions as a way to exert greater control over client behavior. We see substantial discretion as well in the ways that local program managers pursue strategic but perverse organizational maneuvers in response to performance pressures.

Yet a closer look at these examples also serves to underscore the dependence of individual agency on organizational forces. In Foucault's (1980) terms, one might say that as disciplinary power instigates resistance, it also shapes the mentalities of resistance and the terrains that resistance must traverse. The frustrations that influence case managers sanction decisions do not originate in personal life. They are driven by the organization of WT casework, the pressures of competitive performance systems, and the disciplinary frame of reference in which workers are embedded. Likewise, when case managers turn to sanction threats as a way to bring client behavior into line with the organizational imperatives they confront, they are acting in precisely the ways that Lipsky (1980, 1984) emphasized in explaining how bureaucratic processes give rise to discretionary practices of rationing, silencing, and disentitlement. "Burden shifting" case managers who choose to protect their clients swim more directly against the organizational tide, but their small numbers and their stories testify to the difficulty of doing so, the personal costs involved, and the forces arrayed against maintaining such a strategy over the long haul. The perverse organizational responses we observed are facilitated by managerial discretion, to be sure. But as our analysis makes clear, these responses are, for the most part, predictable outcomes driven by the structures and processes in which program managers and frontline workers alike operate.

In this regard, our analysis also raises troubling questions about the NPM and the sharp turn toward performance systems in recent years. We are hardly the first to point out that performance indicators can give rise to perverse organizational responses. But our analysis suggests that such problems may come in a broader variety of forms than previous studies have suggested. Equally important, we find that these problems flow from self-defeating contradictions that are deeply embedded in the NPM—tensions among core principles that are supposed to fit together seamlessly and efficiently. At same time, our analysis also underscores that a narrow focus on the strategic rationality of individuals is likely to understate the scope of perversities engendered by performance pressures. To understand why case managers impose sanctions that they expect to harm their strategic interests, one must look to broader organizational dynamics and consider how performance pressures affect emotions as well as interests.

The NPM suggests to many reformers that they can have their cake and eat it too: centralized control of outcomes and local autonomy; generalization of best practices and diverse solutions tailored to local needs; private provision and public purpose; and competition between regions and collaboration among regions. It sounds too good to be true, and it is. Reformers would be better served by an open acknowledgement that features of the NPM lie in tension with one another and tend to work at cross-purposes. It is not possible to maximize the benefits of organizational diversity and locally tailored problem solving, whereas, at the same time, placing actors in a competitive performance system that promotes conformity and pushes them toward easy rather than challenging paths of innovation. If local diversity and organizational creativity are priorities in a given policy context, performance pressures may need to be eased, at least for a time, to facilitate these goals. And if reformers wish to pursue outcome-focused goals through strong performance systems, they should recognize that their pursuit is likely to come at the cost of some valuable forms of diversity and innovation.

Finally, by employing an organizational lens, our analysis provides some important correctives for the literature on sanctioning in welfare programs. First, sanctioning is more than a response to client behaviors and characteristics and more than an individual action taken by a case manager. Sanctioning is an organized practice, and as a result, the frequency and incidence of sanctions depend on organizational forces. To ignore these factors, as most studies of sanctioning have, is to promote systematic misunderstandings of what sanctions are and how they are distributed.

Second, performance pressures influence sanctioning in ways that can be more deeply perverse than the typical narrative of rational actors sacrificing substantive program goals in favor of their instrumental performance interests. When confronted with poor performance numbers, local service providers respond by sanctioning at higher rates—a reaction that virtually everyone in the WT program expects to *hurt* performance. Far from being a rationally chosen operational improvement, performance-driven sanctioning emerges here as an unintended byproduct of organizational forces. It occurs because frontline workers are under great stress, possess few tools, are pressured to control client behaviors, and get frustrated when client behaviors put their performance at risk.

Last, our analysis highlights a variety of troubling influences on sanctioning patterns that have been systematically ignored in the literature to date. When performance indicators go south, politics matters: ideologically conservative regions respond with higher sanction rates, whereas liberal regions show no change. Cash motives matter too: in a system where high stakes attach to performance pay points, for-profits sanction at systematically higher rates than nonprofits. And power and privilege matter too: when regional performance declines, the clients most disadvantaged by racial and educational factors bear the heaviest burdens of increased sanctioning. In these ways, sanctions are not only products of organization, they are expressions of power, profit, and ideology.

FUNDING

Annie E. Casey Foundation; the University of Kentucky Center for Poverty Research; the Institute for Research on Poverty at the University of Wisconsin; and the Center on Ethnicities, Communities, and Social Policy at Bryn Mawr College.

APPENDIX

Table A1
Variable Definitions, Sources, and Descriptive Statistics for Analyses Presented in Tables 1–3

Variable Name	Definition	Mean	SD	Minimum–Maximum
Individual characteristics (from Table 1)				
Sanction	1 = sanctioned, 0 = otherwise	0.32	0.50	0–1
Gender	1 = male, 0 = female	0.14	0.34	0–1
Age	Client age (in years)	29.69	9.30	18–76
Marital Status	1 = single parent, 0 =otherwise, based on no. of adults in family	0.76	0.43	0–1
Number of Children	Number of children	1.86	1.29	0–13
Citizen	Citizenship status (1 = citizen, 0 = noncitizen)	0.84	0.36	0–1
Age of Youngest Child	Age of youngest child (in years)	5.12	4.95	0–18
Spell Length	Length of TANF Spell (in months)	2.37	1.611	1–14
Race or ethnicity (reference = white, non-Hispanic)				
Black	1 = black, 0 = otherwise	0.43	0.49	0–1
Hispanic	1 = Hispanic, 0 = otherwise	0.28	0.45	0–1
Contextual variables (from Table 1)				
Local Conservatism	See table A2	−0.65	0.220	0–1
For-profit Provider	1 = for-profit provider, 0 = otherwise	0.82	0.38	0–1
Local Unemployment Rate	Unemployment rate in county of client, measured each month (Florida Research and Economic Database)	5.52	1.42	2.51–11.80
Per Capita caseload	Number of TANF adults per 100 county residents (calculated by authors)	0.225	0.112	0.027–0.481
County-level variables (from Tables 2–3)				
Sanction Rate	Percentage of adult caseload that is sanctioned	13.06	3.75	4.47–28.71
Performance Ranking	Monthly change in the average regional performance ranking for the entered employment rate, entered wage rate, and welfare return rate.	−0.01	1.74	−8.33 to 8.33

Continued

Table A1 *(continued)*
Variable Definitions, Sources, and Descriptive Statistics for Analyses Presented in Tables 1–3

Variable Name	Definition	Mean	SD	Minimum–Maximum
Participation Rate	Percentage of the caseload participating in countable work activities (lagged 3 months).	17.31	9.68	0.5–47.4
Average Time on Current Spell	Average number of months on current TANF spell for adult caseload.	4.24	0.68	2.83–7.34
Average Spell Number	Average number of TANF spells of adults on TANF (including current spell)	2.07	0.28	1.50–2.92
Average Number of Children	Average number of children for each TANF adult	1.87	0.10	1.52–2.15
Average Age of Youngest Child	Average age of the youngest child for each TANF adult	5.16	0.47	4.06–6.43
% Single	Percentage of adult caseload that is single.	0.81	0.05	66.96–0.91.35
% Male	Percentage of adult caseload that is male.	1.12	0.03	4.56–0.20.93
% Citizens	Percentage of adult caseload that is a US citizen.	0.95	0.07	59.61–100
% Black	Percentage of adult caseload that is black.	0.40	0.15	7.75–80.74
% Hispanic	Percentage of adult caseload that is Hispanic.	0.12	0.12	0–56.25
Average age of adult clients	Average age of all adults on TANF	29.61	1.06	26.31–31.98
Size of caseload	Number of adult TANF recipients in region			125–9325

Note: Characteristics of the adult caseload are based on administrative records provided by the Department of Children and Families. Data on regional performance ranking, the participation rate, and the size of the adult caseload were obtained from online reports published by the Agency for Workforce Innovation.

Table A2
Construction of Index of Local Political Ideology

Ballot Title	Election Year	Ballot Number	To Amend
Should Two-Thirds Vote be Required for New Constitutionally Imposed State Taxes/Fees?	1996	Constitutional Amendment 1	Art. XI, sec. 7
Fee on Everglades Sugar Production	1996	Constitutional Amendment 4	Art. VII, sec. 9
Responsibility for Paying Costs for Water Pollution Abatement in the Everglades	1996	Constitutional Amendment 5	Art. II, sec. 7
Preservation of the Death Penalty; United States Supreme Court Interpretation of Cruel and Unusual Punishment	1998	Constitutional Amendment 2	Art. I, sec. 17
Additional Homestead Tax Exemption	1998	Constitutional Amendment 3	Art. VII, sec. 6
Public Education of Children	1998	Constitutional Amendment 6	Art. IX, sec. 1
Basic Rights	1998	Constitutional Amendment 9	Art. I, sec. 2
Ballot Access, Public Campaign Financing, and Election Process Revisions	1998	Constitutional Amendment 11	Art. IV, sec. 5a; Art. VI, subsections 1, 2, 5, 7; Art. IX, sec. 4a
Firearms Purchases: Local Option for Criminal History Records Check and Waiting Period	1998	Constitutional Amendment 12	Art. VIII, sec. 5
Florida Transportation Initiative for Statewide High Speed Monorail, Fixed Guideway of Magnetic Levitation System	2000	Constitutional Amendment 1	Art. X, sec. 19
Protect People from the Health Hazards of Second-Hand Tobacco Smoke by Prohibiting Workplace Smoking	2002	Constitutional Amendment 6	Art. X, sec. 20
Voluntary Universal Pre-Kindergarten Education	2002	Constitutional Amendment 8	Art. IX, sec. 1
Florida's Amendment to Reduce Class Size	2002	Constitutional Amendment 9	Art. IX, sec. 1
Animal Cruelty Amendment: Limiting Cruel and Inhumane Confinement of Pigs during Pregnancy	2002	Constitutional Amendment 10	Art. X, sec. 19
Parental Notification of a Minor's Termination of Pregnancy	2004	Constitutional Amendment 1	Art. X, sec. 22
Florida Minimum Wage Amendment	2004	Constitutional Amendment 5	Art. X
The Medical Liability Claimant's Compensation Amendment	2004	Constitutional Amendment 3	Art. I, sec. 26
Authorizes Miami-Dade and Broward County Voters to Approve Slot Machines in Pari-mutuel Facilities	2004	Constitutional Amendment 4	Art. X, sec. 19

Note: The index of local ideology is constructed from data on 18 ideologically relevant constitutional amendments that appeared on Florida's statewide ballot for ratification at some point from 1996 through 2004. The percentage of votes in favor of each amendment is computed for each county. A factor analysis is then conducted using all 18 amendments (thus 18 variables, $N = 67$ counties). The amendments are identified in the table. The final ideology index was aggregated to the regional level for all analyses reported in this paper. For evidence of the validity of this measure, see Fording, Soss and Schram (2007).

REFERENCES

Agency for Workforce Innovation. 2004. *Noncompliance and sanctions policy. Spring 2004*. http://www.floridajobs.org/PDG/TrainingPresentations/WTSanctions 051004.pps (accessed July 24, 2009).

Baumgartner, M. P. 1992. The myth of discretion. In *The uses of discretion*, ed. Keith Hawkins, 129–62. Oxford: Clarendon Press.

Bell, Stephen H., and Larry L. Orr. 2002. Screening (and creaming?) Applicants to job training programs: The AFDC homemaker—Home health aide demonstrations. *Labour Economics* 9:279–301.

Botsko, Christopher, Kathleen Snyder, and Jacob Leos-Urbel. 2001. *Recent changes in Florida welfare and work, child care, and child welfare systems*. Washington, DC: Urban Institute.

Brodkin, Evelyn Z. 1997. Inside the welfare contract. *Social Service Review* 71:1–33.

———. 2007. Bureaucracy redux. *Journal of Public Administration Research and Theory* 17:1–17.

———. 2011. *Policy work: Street-level organizations under new managerialism. Journal of Public Administration Research and Theory* 21(Supp 2).

Considine, Mark. 2003. Governance and competition: The role of nonprofit organizations in the delivery of public services. *Australian Journal of Political Science* 38:63–77.

Dias Johnson, Janice, and Stephen Maynard-Moody. 2006. For-profit welfare: Contracts, conflicts, and the performance paradox. *Journal of Public Administration Research and Theory* 17:189–211.

Ewalt, Jo Ann, and Edward T. Jennings, Jr. 2004. Administration, governance, and policy tools in welfare implementation. *Public Administration Review* 64:449–62.

Feldman, Martha S. 1992. Social limits to discretion. An organizational view. In *The uses of discretion*, ed. Keith Hawkins, 163–84. Oxford: Clarendon Press.

———. 2009. Florida Department of Children and Families. 2009. *OpenClosed TANF Data Report*. http://www.dcf.state.fl.us/ess/agencyforms.shtml#data (accessed June 1, 2009)

Foucault, Michel. 1997. *Discipline and punish: The birth of the prison*. New York, NY: Vintage.

———. 1980. Power and strategies. In *Power/knowledge—Selected interviews and other writings, 1972–1977*, ed. Colin Gordon, 134–45. Brighton, UK: Harvester.

Fording, Richard C., Joe Soss, and Sanford F. Schram. 2007. Devolution, discretion and the impact of local political values on TANF sanctioning. *Social Service Review* 81:285–316.

Gainsborough, Juliet F. 2003. To devolve or not to devolve? Welfare reform in the states. *Policy Studies Journal* 31:603–23.

Gilliom, John. 2001. *Overseers of the poor: Resistance, surveillance, and the limits of privacy*. Chicago, IL: Univ. of Chicago Press.

Goggin, Malcolm, James Lester, Laurence O'Toole, and Ann Bowman. 1990. *Implementation theory and practice: Toward a third generation*. Glenview, IL: Scott Foresman.

Hasenfeld, Yeheskel. 1992. The nature of human service organizations. In *Human services as complex organizations*, ed. Y. Hasenfeld, 3–23. Newbury Park, CA: Sage.

Hasenfeld, Yeheskel, Toorjo Ghose, and Kadyce Larson. 2004. The logic of sanctioning welfare recipients: An empirical assessment. *Social Service Review* 78:304–19.

Heinrich, Carolyn J. 2000. Organizational form and performance: An empirical investigation of nonprofit and for-profit job-training service providers. *Journal of Policy Analysis and Management* 19:233–61.

Holzer, Harry J., and Michael A. Stoll. 2002. *Employer demand for welfare recipients by race. Assessing the new federalism Report 01–07*. Washington, DC: Urban Institute. http://www.urban.org/url.cfm?IDpublications/310423.html (accessed February 3, 2011).

Kalil, Ariel, Kristin S. Seefeldt, and Hui-chen Wang. 2002. Sanctions and material hardship under TANF. *Social Service Review* 76:643–62.

Kettl, Donald F. 2005. *The global public management revolution: A report on the transformation of governance*. Washington, DC: Brookings Institution.

Lipsky, Michael. 1980. *Street-level bureaucracy: Dilemmas of the individual in public services*. New York, NY: Russell Sage Foundation.

Lurie, Irene. 2006. *At the frontlines of the welfare system: A perspective on the decline in welfare caseloads*. Albany, NY: State Univ. of New York Press.

Mashaw, Jerry L. 1983. *Bureaucratic justice*. New Haven, CT: Yale Univ. Press.

Maynard-Moody, Steven, and Michael Musheno. 2003. *Cops, teachers, counselors: Stories from the front lines of public service*. Ann Arbor, MI: Univ. of Michigan Press.

Mead, Lawrence M. 2004. *Government matters: Welfare reform in Wisconsin*. Princeton, NJ: Princeton Univ. Press.

———. 1998. Telling the poor what to do. *The Public Interest* 132:97–112.

Moynihan, Donald P. 2008. *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown Univ. Press.

———. 2006. Managing for results in state government: Evaluating a decade of reform. *Public Administration Review* 77–89.

Pavetti, LaDonna, Michelle Derr, Gretchen Kirby, Robert Wood, and Melissa Clark. 2004. The use of TANF work-oriented sanctions in Illinois. Washington, DC: Mathematica Policy Research, Inc.

Pavetti, LaDonna, Michelle K. Derr, and Heather Hesketh. 2003. *Review of sanction policies and research studies: Final literature review*. Report prepared for Office of the Assistant Secretary for Planning and Evaluation. Washington, DC Mathematica Policy Research.

Radin, Beryl A. 2006. *Challenging the performance movement: Accountability, complexity, and democratic values*. Washington, DC: Georgetown Univ. Press.

Riccucci, Norma. 2005. *How management matters: Street-level bureaucrats and welfare reform*. Washington, DC: Georgetown Univ. Press.

Ridzi, Frank. 2004. Making TANF work: Organizational restructuring, staff buy-in, and performance monitoring in local implementation. *Journal of Sociology and Social Welfare* 31:2.

Schram, Sanford F., Joe Soss, Richard C. Fording, and Linda Houser. 2009. Deciding to discipline: Race, choice, and punishment at the frontlines of welfare reform. *American Sociological Review* 74:398–422.

Talbot, Colin. 2005. Performance management. In *The Oxford handbook of public management*, eds. E. Ferlie, L. E. Lynn, Jr., and C. Pollitt, 491–517. Oxford: Oxford Univ. Press.

Watkins-Hayes, Celeste. 2009. *The new welfare bureaucrats: Situated bureaucrats and entanglements of race, class, and welfare*. Chicago, IL: Univ. of Chicago Press.

———. 2011. Race, respect, and red tape: Inside the black box of racially representative bureaucracies. *Journal of Public Administration Research and Theory* 21(Supp 2):i229–i47

Weisbrod, Burton A. 1998. To profit or not to profit: The commercial transformation of the nonprofit sector. New York, NY: Cambridge Univ. Press.

Weissert, Carol, ed. 2000. *Learning from leaders: Welfare reform politics and policy in five midwestern states*. Albany, NY: Rockefeller Institute.

Wu, Chi-Fang, Maria Cancian, Daniel R. Meyer, and Geoffrey Wallace. 2006. How do welfare sanctions work? *Social Service Review* 30:33–51.