An Empirical Investigation of Gaming Responses to Explicit Performance Incentives

Pascal Courty

London Business School

Gerald Marschke

State University of New York, Albany [1,2]

August 29, 2000

Abstract: This paper studies a particular kind of gaming responses to explicit incentives in a large government organization. The gaming responses we consider occur when agents strategically report their performance outcomes to maximize their awards. An important contribution of this work is to examine whether this behavior diverts resources (*e.g.* agents' time) from productive activities or whether it simply reflects an accounting phenomenon. We evaluate the efficiency impact of the behavior we identify and find that it has a negative impact on the true goal of the organization. (*JEL* J33, L14)

[1]Comments welcome at Pascal Courty, Department of Economics, London Business School, Regents' Park, NW1 4SA, London, UK, pcourty@london.edu and Gerald Marschke, Department of Economics and Department of Public Administration & Policy, BA-110, SUNY-Albany, Albany, NY 12222, marschke@csc.albany.edu.

[2]Courty and Marschke acknowledge support from NSF-SBR-91-31143 to J. Heckman. Courty acknowledges support from the Sloan Foundation and The University of Chicago. Marschke acknowledges support from a grant from the Upjohn Institute and a grant from the American Bar Foundation. We are grateful for the support and encouragement of James Heckman. We thank Michael Gibbs, Robert LaLonde, Canice Prendergast, Sherwin Rosen, Jeffrey Smith and seminar participants at Chicago, Rochester, Arizona, IESE, and Bristol for helpful suggestions. All errors are exclusively the responsibility of the authors.

1 Introduction

When performance awards depend non-linearly on performance outcomes, agents have an incentive to manipulate the timing of their performance. Under bonus-based contracts, for example, agents may time their performance so that they just meet the numerical standard to receive their bonus. Under this timing strategy, agents postpone excess performance above the standard from good to bad years thereby increasing their chances to receive bonuses in bad years. Although these timing responses to non-linear contracts have received some attention in the organization literature, there is no direct evidence that this behavior creates a welfare loss. (See Gibbons (1997) and Prendergast (1999) for a survey.) The main goal of this paper is to examine whether this behavior reflects a misallocation of real resources or simply an accounting phenomenon.

The distinction between timing responses based on an efficiency criterion plays a key role in this paper. Timing strategies that create no welfare loss while increasing the agents' chances of earning bonuses are accounting manipulations. Accounting responses do not create any welfare loss because the organization can neutralize this behavior by appropriately discounting the performance award. Alternatively, timing strategies may lead agents to "invest in the wrong tasks" or to "game" the performance incentive system (Holmstrom and Milgrom, 1991 and Baker, 1992).[1] Timing strategies are evidence of gaming if they do not only imply some kind of accounting manipulation, but also a costly misallocation of resources.

We investigate this distinction between accounting and gaming responses in a large federal job training program for the disadvantaged, created by the Job Training Partnership Act (JTPA) of 1982. JTPA is one of the first large-scale implementations of financial performance incentives in a federal bureaucracy. The JTPA incentive system rewards training agencies on the basis of their trainees' labor market achievements on the date they graduate from training. The incentive system leaves training agencies some discretion over the timing of the graduation date. The central focus of this paper is to study

[1] We take in this work a wide interpretation of the multi-tasking framework and consider the timing strategy as one of the agent's tasks even though it has no value to the principal.

how training centers use this discretion.

Some of the timing strategies presented here have already been identified in our previous work (Courty and Marschke, 1997). This paper builds on that work in three important ways. First, this paper offers an analytical description of the training center's optimal graduation strategy that generates a much wider set of implications than were considered previously. Second, we present formal tests of the model's implications. Third and most importantly, this paper investigate some of the welfare implications of timing strategies.

The paper develops as follows. The following section reviews the organization literature on gaming responses to explicit incentives. Section 3 describes the JTPA organization, its performance incentive system, and our data. Section 4 shows that the program administrators maximize their performance awards by timing optimally the graduation date of program participants. Section 5 studies the efficiency of these timing strategies and Section 6 summarizes our findings.

2 Incentive Theory and Gaming Evidence

Bonus-based compensation schemes award bonuses when performance outcomes exceed preset numerical standards. Such schemes are common in some occupations: CEOs, sales people, and piece workers receive bonuses when firm earnings, yearly sales, and piece rates exceed predetermined targets, quotas or thresholds (Murphy, 1998). An important regularity that is emerging from the empirical literature on incentives shows that nonlinear incentive contracts often encourage agents to manipulate the timing of their performance.

In an early contribution to this literature, Healy (1985) documents that managers who are compensated for meeting annual income thresholds use their discretion over the timing of income reporting to smooth their compensation across accounting years. More recent works report similar timing responses to threshold effects in other settings. For example, Asch (1990) showed that navy recruiters who receive awards for meeting year-end recruitment quotas respond by reallocating their work efforts over the year. Similarly, Oyer (1998) showed that there is more variability in firms' sales at the end of the fiscal years—when sales persons' bonuses are computed—than in the middle. The empirical evidence on threshold-motivated timing responses to incentives is surveyed in Gibbons and Prendergast.[2]

Four components common to many incentive systems create these timing responses. First, the incentive award is a non-linear function of the agent's performance. Second, the performance outcomes are aggregated over fixed periods and the agent is rewarded at the end of each period on the basis of her cumulated performance. Third, the performance outcomes vary for random reasons that are outside the agent's control. Finally, the agent is able to choose the period in which performance is reported. Although the argument applies generally to non-linear awards, the rationale behind these timing responses is best illustrate under a bonus-based award.

Consider a simple environment where random fluctuations imply that the agent would succeed in winning a bonus in good years but fail in bad years if she reported performance outcomes without delay. Now, consider the following strategy that takes advantage of the agent's discretion over performance reporting. In bad years, the agent delays performance below the standard until the next bonus year, while in good years the agent delays positive performance above the standard. In very bad years, it may be optimal to 'take a bath', that is, to report bad performances and carry over good performances until the next bonus year. Under this timing strategy, the agent is more likely to receive a bonus in bad years without compromising the chances of winning bonuses in good years.

In exposing this timing behavior, the literature shows the variety and creativity of agent responses to incentives but offers no direct evidence that such behavior creates a welfare loss. In fact, the literature does not make an explicit distinction between accounting and gaming responses as defined in the Introduction. Based on this distinction, a review of the timing literature suggests that agents manipulate performance accounting

[2]There is also a literature studying timing responses by tax-payers to minimize their tax bill (e.g. Dickert-Conlin and Chandra,1999, and Goolsbee 2000). Brown, Harlow and Starks (1996) and Chevalier and Ellison (1997) show that fund managers modify their investment strategies over the year to maximize investment flows into their funds. These studies demonstrate timing responses to incentive schemes. Non-timing responses have also been documented. For example, Drago and Garvey (1997) offer evidence that individual-based performance incentives discourage coworkers from helping one another.

but is inconclusive regarding whether this manipulation is gaming. For example, the managers in Healy's study may not have to consume resources in figuring out at the end of the accounting year how to optimally report their financial performance outcomes. Similarly, the military recruiters in Asch's study as well as the salespeople in Oyer's study may vary their effort supply over the contract year as a result of the incentives they face, but this by itself is not evidence of inefficiency. This work is the first attempt to estimate empirically the costs of timing activities in performance incentives.

Discovering whether timing strategies are inefficient has important economic implications. Evidence of welfare losses would contribute to a general argument, found in the organization literature, attributing the scarcity of explicit incentives to the difficulty of measuring performance outcomes. (See Jensen and Meckling (1992) for a review.) According to this argument, principals can specify explicit performance measures only approximately, which results in agents gaming the system by optimizing with respect to the actual measures instead of the intended unmeasurable objective (Holmstrom and Milgrom and Baker). Despite the popularity among organization theorists of the gaming argument as an explanation for the scarcity of explicit incentives, there is little formal evidence in its support. In fact, the gaming argument relies exclusively on anecdotes accumulated over the years by organizational behaviorists (e.g. Kerr (1975) and Lawler (1982)). Although these anecdotes suggest that explicit incentives do not always work as intended, they do not demonstrate that these unintended responses are evidence of inefficiency.

The set of threshold-driven timing responses sketched above seem to be a good candidate to conduct a formal test for costs of the unintended responses. We propose to conduct this test in the JTPA organization for two reasons. First, the program administrators in JTPA face a similar incentive environment as those studied in previous empirical work. Meeting a fixed numerical performance standard plays a key role in the JTPA award functions. Actually, several award functions in our sample are identical to a bonus-based formula. Second, a recent evaluation of JTPA has generated a wealth of data on its functioning, including separate accounts, by program administrators and by participants, of the training outcomes. Our ability to observe agent behavior that is normally hidden from the organization, at least in the short-run, makes this study especially compelling. In addition, these data are sufficiently rich to generate measures of agent productivity. In contrast with previous timing studies, these measures will allow us to conduct the efficiency study that is required to test the gaming hypothesis.

Our strategy will be to identify some timing responses and to sign their welfare impacts. Unfortunately, our data do not allow us to conduct a full-scale welfare analysis. This study, however, presents findings that should be taken into account in a more complete evaluation of the overall efficiency of the JTPA incentive system.

3 JTPA: Organization, Incentive System and Data

This section starts with a few summary facts about job training under JTPA, then describes the JTPA incentive system emphasizing those features that are most relevant for this study and finally introduces the three data sources that will be used in this paper. The Job Training Partnership Act of 1982 created what is presently the largest federal employment and training program serving the disadvantaged. Its current annual budget is approximately $4 billion and it serves nearly one million people annually.[3] The JTPA bureaucracy is unusual for three reasons. First, JTPA is highly decentralized: job training is carried out by more than 620 semi-autonomous sub-state training agencies. Second, the Act gives these training agencies significant discretion over who is admitted to the program and how training is conducted. Third, and most important for this study, instead of a rigid, comprehensive set of rules to regulate bureaucratic *conduct*, the federal government uses a loose set of financially-backed performance incentives to influence *outcomes*.

[3]For a detailed description of JTPA, see Johnston (1987). For detailed descriptions of its incentive system see Barnow (1992), Courty and Marschke (2000), and Dickinson *et al.*, (1988). For empirical analyses of the effects of incentives on the population selected to receive training see Heckman, Smith and Taber (1996), Anderson, Burkhauser, and Raymond (1993), Cragg (1997) and on the service provided see Marschke (2000).

3.1 The Incentive System

In this section, we describe the main features of the incentive system which will be the focus of our empirical work. Congress intended the performance incentives to measure the training agency's success in developing participants' human capital (Job Training Partnership Act, Sec. 106(a)).[4] The Act gave the U.S. Department of Labor (DOL) the responsibility to develop a workable set of performance measures based on the Act's mandate. The DOL chose a set of short-term labor-market measures based on an enrollee's employment status, wage, earnings, and/or number of weeks worked at the end of job training. Table A.1 in the appendix defines the federal measures.

Along with selecting labor market measures, the DOL also decided that the labor market outcomes would be measured on the date participants graduate from the program.[5] The graduation date is the date the training agency officially closes an enrollee's case and removes her from its rolls. To receive additional training after the graduation date, the enrollee must be re-enrolled. Note that the same date corresponds to two concepts: *graduation*, that is, termination of the enrollment period, and *reporting* of the labor market outcome.[6] Training agencies discretion over the graduation date is limited by the following rule: the JTPA regulations require training agencies to graduate a participant—at which time her labor market outcomes are recorded—within 90 days of her last day of training.

The individual states administer the incentive system. The JTPA fiscal year lasts from July 1 to June 30 of the next calendar year. At the end of each fiscal year, or *program* year, the state rewards (or sanctions) their training agencies on the basis of their year-end performance outcomes. In aggregating performance over a fixed period of 12 months, the JTPA incentive system is thus similar to many incentive systems. The award augments the training agency's budget in the following year. Averaging over all JTPA training

[4]It should be noted that human capital development is not the exclusive goal of JTPA. The Act, the Act's amendments, and the U.S. Department of Labor in its role as interpreter of the Act mention other goals such as equity and special service to individuals in the bottom of the income distribution. See Heckman and Smith (1995) for a discussion of these goals and how they are expressed through the JTPA performance incentives.

[5]In the time period for our study, 1987-89, the date of performance measurement corresponds to the graduation date.

[6]We thank Canice Prendergast for suggesting this terminology.

agencies, the award is the source of approximately seven percent of the operating budget.

3.2 The Data

This study relies on three data sources. The first two data sources are an administrative data source and a participant data source which were collected for the DOL-commissioned National JTPA Study (NJS). The NJS was an experimental study of the effectiveness of JTPA conducted between 1987 and 1989. The study was conducted using a classical experiment methodology according to which JTPA applicants were randomized into treatment and control groups. Members of the treatment group received training and members of the control group were denied JTPA training services for 18 months. Sixteen of the organization's roughly 620 job training agencies participated in the NJS.[7] The administrative data include enrollee-level information on training activities, enrollment and graduation dates, and the employment-at-graduation performance outcome. The participant data were produced from a questionnaire administered to participants at the time of their application and again 18 months later. As a result, this second data source includes detailed information on individuals' participation in job training, schooling, and job search and on employment during the 18 month period following random assignment.

The third data source describes the JTPA incentive policies. We obtained these policies from the states' governors offices.[8] We briefly summarize the main characteristics of the incentive policies that will be used in the paper. Although each state uses similar sets of performance measures, the policies show that the award formulas differ somewhat across states. To illustrate a stylized incentive contract, let S^k be the training agency's performance *outcome* on *measure* k, and $\bar{S}^k$ be the numerical *standard* for the same measure. The simplest type of policy would give an award ψ^k for each measure k only if the reported outcome exceeds its standard. Thus, the program year's incentive award is,

$$A = \sum_k \psi^k \, 1\{S^k > \bar{S}^k\}, \tag{1}$$

[7] See Doolittle and Traeger (1990) for a description of the implementation of the NJS. See Orr *et al.* (1994) for a detailed description of the results of the NJS.

[8] We collected these data while research associates for the Center for Social Policy Evaluation, at the University of Chicago. A summary description of this data can be found in Courty and Marschke, 2000.

where $1\{\cdot\}$ equals one if the condition in brackets is true. Because the Act leaves the design of the incentive policies to the individual states, there is much variation in the precise functional form of the awards. In some states, training agencies receive additional awards for exceeding standards, with the performance-award relationship displaying many sorts of non-linearities, such as piece-wise linear forms and nested qualifying conditions.

What matters for this study is that the Act requires only that awards be contingent on achieving standards. As a consequence, the major portion of the potential award across all training agencies in our study is paid out for simply meeting the standard. For our analysis, we will assume that each training agency faces functional form (1) because it is the simplest formulation allowing us to test for incentive responses common to all incentive policies.

4 Timing Activities in JTPA

In this section, we study how training agencies time the graduation date. We start by investigating the optimal graduation strategy before the end of the program year assuming that the award depends only on the employment performance measure. We then generalize this simple graduation strategy. The first two subsections develop and test econometric models based on the argument sketched out in Courty and Marschke (1997).[9] The final subsection extends the argument of Courty and Marschke to examine the effects of features of the incentive system that they did not consider.

4.1 Graduation Timing Before the Program Year-End

Because labor market outcomes vary over time naturally on their own, training agencies have an incentive to manipulate the reporting date for the employment outcomes. At the end of an enrollee's training, training agencies face a decision: to graduate the enrollee and report her labor market outcomes, or to postpone graduation in hopes that the outcome improves. The optimal graduation strategy leads the training agency to graduate enrollees

[9]We borrowed Figures 1, 2 and 3 from Courty and Marschke to concisely convey the main intuition of their argument.

who are employed within the 90 day period following training either on the last day of training or on the first day of employment, whichever occurs first, and all others on the 90th day following training end.

We introduce some notation to expedite the discussion. Let g_i and f_i be enrollee i's graduation date and training end date, respectively. Figure 1 presents the distribution of $g_i - f_i$, the number of days that pass between the end of training and graduation. This distribution is plotted separately for persons who report being employed and unemployed on the last day of training. First note the wide dispersion of $g_i - f_i$. The fact that many enrollees are reporting that their training ends after the graduation date (negative values of $g_i - f_i$), we believe, indicate considerable error in the measurement of the training end date.[10] Even so, among enrollees who are employed on the last day of training (represented by the dashed line in Figure 1), the modal enrollee is graduated at the first possible opportunity, that is, at the end of training. For the modal enrollee in that group, $g_i = f_i$. Among enrollees who report no employment at training end (represented by the solid line in Figure 1), enrollees are not systematically graduated on the last day of training. The presence of a pronounced spike in graduations for employed at training end, and not for persons who are unemployed is consistent with the optimal graduation rule.

Figure 2 shows the same distribution by whether the enrollee is graduated employed. The solid line in Figure 2 shows that the modal person who is unemployed at graduation is graduated at the last possible moment, about ninety days after training ends. The dotted line in Figure 2 shows the absence of a pronounced spike in graduations at the end of the ninety day window for enrollees who graduate employed. Instead, among those enrollees who graduate employed, graduations are highest in the days immediately following the training end, tapering off through the end of the ninety day window, a pattern consistent

[10]We believe that there is considerable measurement error in the enrollee-reported training end date for two reasons. First, because for many enrollees a year or more has passed between the last day of JTPA training and the NJS follow-up survey, enrollees imprecisely recall the training end date. Second, the survey inquires about all past training spells. While the survey seeks information about each past spell individually, it does not ask the enrollee to indicate whether the spell occurred in a JTPA program. Because JTPA spells are not identified, some spells that we treat as JTPA spells are not JTPA training spells. Where an enrollee reports multiple training spells we consider the spell with the end date closest to the graduation date for the comparison.

with the fact that many enrollees who were initially unemployed found a job within 90 days of the end of training.

We now test for the optimal graduation strategy more formally. Let e_i be the earliest date following f_i that individual i is employed, and u_i be a dummy indicating whether enrollee i failed to find a job within ninety days after f_i. Using this notation, the optimal graduation strategy states that training agencies graduate (1) enrollees who are employed within the ninety day period on their first employment after training ends (that is, $g_i = e_i$ if $u_i = 0$) and (2) enrollees who were never employed within the 90 days period on the 90th day following the end of training (that is, $g_i = f_i + 90$ if $u_i = 1$). Thus, we have

$$g_i = (1 - u_i)e_i + u_i(f_i + 90). \tag{2}$$

To test this identity, we assume that participants report their training end date and employment date with the same error, and that the training agency reports the graduation date with a random error that is stochastically independent of the participant's error. The observed variables are u_i, $\tilde{g}_i = g_i + \eta_{g_i}$, $\tilde{f}_i = f_i + \eta_i$, and $\tilde{e}_i = e_i + \eta_i$, with η_{g_i} and η_i assumed to be stochastically independent. Taking the graduation date as a reference point, (2) can be rewritten in terms of the observed variables,

$$\tilde{g}_i - \tilde{f}_i = (1 - u_i)(\tilde{e}_i - \tilde{f}_i) + u_i 90 + \nu_i, \tag{3}$$

where $\nu_i = \eta_i - \eta_{g_i}$. To test the identity (3) we add coefficients and estimate

$$\tilde{g}_i - \tilde{f}_i = \beta_1(1 - u_i)(\tilde{e}_i - \tilde{f}_i) + \beta_2 u_i + \nu_i. \tag{4}$$

Because the residual is stochastically independent of the independent variables, we estimate (4) using least squares. Under the assumption that the error term is normally distributed, we can test whether the coefficients corresponding to $(1 - u_i)(\tilde{e}_i - \tilde{f}_i)$ and u_i, that is, β_1 and β_2, equal 1 and 90, respectively.

Model I, Panel A, of Table 1 presents a simple comparison of the mean delay for enrollees who are employed sometime between training end and the end of the ninety day period (enrollees for whom $u = 0$) and enrollees who never become employed ($u = 1$).[11] (Model I is a regression of graduation delay on the variables $(1 - u)$ and u, without an intercept.) Model I shows that participants who are never employed during the ninety days following training end graduated 67 days later than those participants who obtain employment sometime during the ninety day period. Panel C of Table 1 shows the p value of the test of significance of this difference to be almost zero.

Model II in Table 1, reports the estimates of equation (4). The estimated coefficient on $(1 - u_i)(\tilde{e}_i - \tilde{f}_i)$, at .99, to be close and statistically identical to one (the p value of the test of identity to one is .888). The coefficient on u_i equals 101.8 and is significant. Thus, on average, training agencies wait 102 days to graduate those enrollees who remain unemployed after completing their training. The coefficient on u_i is significantly different from 90, however. The observation that training agencies wait a little longer than 90 days to graduate some unemployed enrollees suggests that the 90 days constraint is not always perfectly enforced.

4.2 Graduation Timing Toward Program Year-End

We generalize the simple graduation strategy presented above to take into account the training agency's discretion over the program year in which it graduates participants. Consider a stylized two program year incentive system where the training agency receives a fixed payment if the yearly labor market-based performance outcome exceeds a fixed performance standard. The training agency does not know its final aggregate performance outcome until the end of the program year because the labor market outcomes depend upon random factors, such as the state of the local economy, which are outside its control. Because of the graduation strategy described above, the training agency reaches the end

[11] The NJS data files contain 6444 adults with valid training agency-supplied enrollment and graduation dates who graduated in program years 1987 through 1989. Nevertheless, many of these participants failed to report in the NJS's participant survey having experienced a job training spell. Moreover, many who reported training spells supplied invalid beginning and ending dates. For some of these spells, however, we were able to impute dates. The subsample used in our graduation delay study described here contains, after imputing some training spell dates, 2327 persons. An appendix describing in more detail the construction of this subsample, including the imputation procedure, is available upon request.

of the year with an inventory of enrollees who have finished training within the previous 90 days but are unemployed. At the end of the first program year, the training agency chooses how many from this inventory to graduate in the present program year, the remainder to be graduated in the following program year. Assume there are n such persons, of whom n_1 will be graduated in the first program year and $n_2 = n - n_1$ in the next one. The training agency chooses n_1 to maximize the present value of the sum of the two awards.

The optimal graduation strategy on the last day of the first program year depends on the difference between the performance outcome and the standard as the last day arrives. Let $N = N_e + N_u$ be the number of persons who were graduated during the year (excluding the year's last day), where N_e and N_u are the numbers of such persons graduated employed and unemployed, respectively. Let $\bar{S}$ be the performance standard. Three cases can be distinguished (see Figure 3). In case one, on the last day of the year the cumulative performance outcome exceeds the standard by so much that the training agency can graduate all unemployed enrollees. In case one, because $\frac{N_e}{N+n} \geq \bar{S}$, $n_1 = n$. In case two, the cumulative performance outcome exceeds the standard, but not by much. In case two, because graduating all unemployed enrollees would push the outcome below the standard, it pays the training agency to graduate persons from its inventory only until the standard is bound. That is, the training agency chooses n_1 such that $\frac{N_e}{N+n_1} = \bar{S}$. Rearranging yields

$$n_1 = \frac{N_e}{\bar{S}} - N. \tag{5}$$

Equation (5) implies that n_1 lies between 0 and n, approaching zero when the training agency just meets the standard and n when the training agency outperforms the standard by n/N or more. In case three, the training agency fails to meet the standard at the end of the year, ($\frac{N_e}{N} < \bar{S}$). In this case, because it cannot win an award this year, the training agency 'takes a bath', graduating all n persons from its inventory to maximize the probability of an award next year.

Because the risks of not meeting a standard are substantial, the return on the strategy described above is likely to be high. For example, according to Department of Labor data on the entire JTPA system, in 1988, 13.9 percent of training agencies did not achieve the standards for the employment rate at graduation. If the above strategy plays an important role, training agencies should graduate a disproportionate fraction of the enrollees in June.

Figure 4 shows the aggregate monthly graduation counts for the set of sixteen experimental sites after correcting for the uneven sampling introduced by the NJS. Notice that a large fraction of the year's graduations occurs in the month of June.[12] In examining training agencies individually, we found a program year end pile-up in at least one program year in fourteen of the sixteen sites. The program year end pile-up is not explained by the enrollment patterns of the JTPA enrollees, i.e. as Figure 3 shows, the graduation pattern does not appear to be a rightward translation of the enrollment series. Although many JTPA enrollees enroll in classroom-training services offered through community colleges whose school years ended in May or June, this academic cycle does not explain the program year end pile-up. Figure 5 shows the presence of pile-ups even for those enrollees who do not participate in classroom training.

The two-period model outlined above suggests the following modification of (4).

$$\begin{aligned}\tilde{g}_i - \tilde{f}_i &= \beta_1(1-u_i)(\tilde{e}_i - \tilde{f}_i) + \beta_3 u_i(1-j_i) + \beta_5 u_i j_i LOW_i \\ &\quad + \beta_6 u_i j_i MED_i + \beta_7 u_i j_i HIGH_i + \nu_i, \end{aligned} \tag{6}$$

where j_i is a dummy, equal to one if June 30th falls within the 90 day period following training and zero otherwise, and LOW_i, MED_i and $HIGH_i$ are three dummies, respectively equal to one if the performance outcome at the end of the program year lies in the low, medium and high regions defined by the horizontal axis in Figure 3, and zero otherwise.

For persons who are not employed at training end, the model predicts the following:

[12] Training agencies enroll persons into their programs continuously, and year-round. Our data only contain the enrollees who were also part of the NJS. For most training agencies in the NJS, random assignments began in 1987, continued through 1988, and ceased in 1989. Thus, only in 1988 were the majority of the program's enrollments at these sites contained in the experimental group. In program year 1987, typically only persons who enrolled in the latter part of the year were experimental participants. In program year 1989, typically only persons enrolled in the first part of the year were experimental participants. Because the NJS truncates 1987 and 1989, the June pile-ups produced by the raw data in these years are likely to be biased estimates of the true pile-ups. Figure 4 shows graduation counts in which persons enrolled in program years 1986, 1987, and 1989 are re-weighted so that the weighted enrollment patterns in those years are identical to the enrollment pattern in 1988.

1. Unemployed enrollees whose training ends in the last three months of the program year are less likely to be delayed 90 days. As a consequence, we should find that the length of delay is longer on average for persons who finish training between July and March than for persons who finish between April and June. We test whether the length of delay differs in this way by constraining $\beta_5 = \beta_6 = \beta_7 = \beta$, estimating (6), and testing whether $\beta_3 > \beta$.

2. Among unemployed persons whose training ends within 90 days of June 30th, enrollees in LOW_i and $HIGH_i$ categories should be delayed for a shorter period on average than individuals in MED_i. In terms of (6), this corresponds to testing whether $\beta_6 > \beta_5$ and $\beta_6 > \beta_7$. In addition, individuals in LOW_i and $HIGH_i$ should be delayed approximately the same length of time, that is, β_5 should be equal to β_7.

Models III and IV in Table 1 test these predictions and provide evidence consistent with the two-period model of the graduation decision. Model III reruns Model II, splitting enrollees who are not employed throughout the grace period (enrollees for whom $u = 1$) into two groups: one group contains enrollees whose 90 day grace period includes June 30, $j = 1$, and the other group contains enrollees whose 90 day grace period does not include June 30, $j = 0$. As predicted, we find the estimate of graduation delay is greater for the average person who finished training between July and March, compared to the average person who finished between April and June; the point estimates of graduation delay past the end of training is 113 days for the former group, and 79 days for the latter group. This difference is statistically different (the p value of the test of equality is .02). In addition, the R^2 rises slightly, from .150 to .153.

To test the remaining implications of the theoretical model, we created a subsample based on the graduation decisions for the subsample of training agencies in our data for which we had reliable estimates of year-end performance.[13] The model predicts that

[13] Although training agencies provided us with the complete incentives in place during the years 1987 through 1989, they were generally unable or unwilling to provide reliable estimates of their year end performance. Creating reliable estimates of year-end performance from our data is challenging. Because of the design of the NJS, a training agency's experimental population generally represented only a portion of the year's participants. To develop reliable estimates of year-end performance, we omitted training

training agencies will delay graduating the average person whose training ends in MED for a longer period of time than the average person whose training ends in either HIGH or LOW. The estimates of Model IV are consistent with this prediction. That is, the estimated coefficient for the variable $u\ j\ MED$ is greater than the estimated coefficients for the variables $u\ j\ HIGH$ and $u\ j\ LOW$. These differences are statistically significant (the p values for two-tail tests of equality are .030 and .024, respectively). Model IV shows that training agencies apparently exploit the ninety day grace period fully: the coefficient estimates for $(1-u)\ (e-f)$ and $u\ (1-j)$ are very close and statistically identical to one and ninety, respectively.

To summarize, Table 1 shows that a training agency (1) delays graduating idle, unemployed enrollees longer than idle, employed ones, (2) graduates idle, unemployed enrollees sooner if they finish at the end of the year than if they finish at the beginning, and (3) among idle, unemployed enrollees who finish training in the last three months of the year, graduates them sooner if the training agency is doing either very well or poorly relative to the employment standard. These findings are consistent with the two period graduation model.

Another way to display the optimal graduation strategy in the two-period model is by plotting the movement of the employment outcome versus the standard over the course of the program year. The model predicts that under the simplest incentive policy that pays lump sum awards for meeting standards (as in equation (1)), the year end employment outcome of training agencies in the MED or HIGH states must dip sharply downwards toward the standard. For more general incentive policies, the model predicts that the employment outcome should dip toward some level above the standard until the marginal values of unemployed participants is equalized across adjacent years.[14] Thus, the model

agencies from the estimation for which we did not have a substantial portion of the total year's participants. The estimates of Model IV are based on graduation decisions in training agencies for which 70 percent or more of participants graduated during the program year were represented among the NJS participants. For this reason, Model IV is estimated on a subsample of the data used to estimate Models II and III. We also ran Model IV on samples created using training agencies for which we had 60, 50, 40 and 30 percent of the year's population of graduates. We found that these samples generally produced statistically insignificant relationships, presumably because we were admitting sets of observations with ever more noisy measures of year-end excess performance.

[14] By a more general incentive policy we mean one that pays some award for outcomes in excess of the

implies that the *average* performance outcome for training agencies in MED and HIGH states will drop sharply at the year's end toward the standard but not necessarily touch it. Figure 6 is consistent with this prediction.[15] Figure 6 shows a sharp downward movement in the employment outcome toward the standard, consistent with the practice of holding unemployed enrollees in an inventory, drawing down the inventory only when achievement of the standard is assured, and then only until either the overall outcome attains the standard, or the inventory is exhausted. Such a graduation strategy maximizes non-linear intertemporal awards.

Because the optimal strategy implies that individuals graduated in June do not perform as well as the individuals graduated in any other month, we present Table 2, which examines five of the core federal performance measures for June and non-June graduates for all program years.[16] June graduates show significantly lower outcomes on all performance measures for all years. Another way to illustrate the prediction that June graduates should perform worse than non-June graduates is to reproduce Figure 4 separately by whether the enrollee is employed at graduation. Figure 7 shows that only the non-employed produce the June phenomenon.

Let us consider the impact of the graduation strategy on the performance outcome and the performance award. One measure of the private gains from delaying graduating unemployed enrollees is the difference between the fraction of enrollees who are employed on the day they finish their training and the fraction employed on the day the training agency officially reports them. We find that the overall employment rate at graduation increases by 11.3 percentage points between these dates, from 47.0 percent to 58.3 percent. Stated another way, training agencies in this study would produce an employment rate outcome 20 percent lower if they were required to graduate enrollees (and report their performance outcomes) on the date they actually finish training.

standard.

[15] This figure tracks the employment outcome and standard during the program year 1988 for the subset of NJS training agencies whose estimated year-end performance outcomes placed them in either the MED or HIGH states at the end of program year 1988.

[16] We have also conducted this comparison for each of the program years 1987, 1988 and 1989, separately. The performance outcomes in each of these years show the same qualitative differences as in Table 2 .

Finally, one could imagine that training centers might also delay graduating some enrollees who are employed at training end. For example, towards year-end, when training centers are in the HIGH or LOW states, they might delay graduating employed enrollees who had finished their training. While they may sometimes do this, this strategy is likely to be second-order. Delaying graduating employed enrollees would be second-order because until the year-end outcome is assured, a policy of postponing the graduation of employed enrollees risks losing the employment, both because enrollees experience high-turnover, and because caseworkers cannot easily verify the employment status of enrollees who are not actively training.[17]

4.3 Additional Performance Incentives

We now generalize the graduation strategy presented in Section 4.1 to take into account performance measures based on wages and earnings, the other significant class of JTPA performance measures.[18] These other measures may influence the graduation decision although their impact is likely to be small for two reasons. First, the employment-based performance measure plays a disproportionate role in the determination of the award. Second, there is little within-person wage variation during the 90 day period following training end.

Nonetheless, taking into account the wage measures, training agencies may choose to graduate employed enrollees who have little chance of experiencing a wage increase, but wait on those employed enrollees who have a high likelihood of experiencing a wage increase. The training agency does not wait on all enrollees because by doing so it might lose credit for an employment. This risk is significant because approximately one quarter of the enrolled who were employed on their graduation date were not employed at the

[17] An implication of such a strategy is another pile-up of graduations in the first month of a new program year. Figure 4 shows no graduation pile-up in the first and second month of the program year.

[18] Some states in our sample reward training agencies for employments three months after the graduation date. That is, they receive awards if among their graduates the rate of employment on the 90th day after graduation exceeds a numerical threshold. Like the graduation-date based measures, these follow-up measures may influence the graduation decision. Nevertheless, because it is more difficult for the training agency to influence whether a person is employed three months after training ends, follow-up measures should have a smaller influence on the graduation decision.

same job three months later. The refined strategy which takes the wage measure into account implies that some employed enrollees should be graduated later than is predicted under the simple strategy presented previously. This refined strategy is more difficult to identify statistically because the econometrician does not observe many of the variables that are observed by the training agency that determine the likelihood that a particular individual will experience a higher wage in a short horizon. To simplify the analysis, we focus on the enrollees who were employed at the end of training and experienced a second employment spell under a new employer before the close of the 90 day window.[19]

Taking into account the wage performance measure, the optimal graduation strategy implies that those individuals who experience a wage increase on their second employment spell after training ends are more likely to be graduated during that second employment spell while those who experience a wage decrease are more likely to be graduated during the first employment spell. Table 3 presents a contingency table for these two subsets where the two contingent variables are whether the enrollee's graduation was postponed beyond the start of the secondary employment spell and whether the enrollee experiences a wage increase. A secondary employment spell is a spell which began after the month training ends, but before the end of the 90 day grace period.

The upper panel of Table 3 shows that training agencies are more likely to postpone graduating the enrollees who experience a wage increase within the 90 days following training end (respectively 50% and 64% of the enrollees who experience a lower and higher wage are delayed). While persons who experience wage increases show a greater incidence of graduation delay, we cannot reject the alternative hypothesis of no association between the two variables (The P-value from a chi-square test of the hypothesis of no association is .147).

[19]For the analysis reported here, we created a subsample from the sample of 2511 used in the analysis of Section 4.1 and 4.2, by selecting persons who (1) graduated in an incentive regime containing the average wage at graduation measure, (2) had at least one employment spell ongoing at training end and at least one other employment spell beginning before the end of the 90 day grace period, and (3) had valid wage measures for each reported employment spell. The subsample created using these criteria contained only 133 persons. Comparatively few observations are generating the results observed here, because only 8 percent of our sample experience a second employment spell within the 90 day period. An appendix describing in detail the construction of this subsample is available upon request.

The optimal graduation strategy also predicts that those participants who are graduated after the start of their secondary employment spell should experience higher wage offers. The lower panel of Table 3 tests for a relationship between the size of the wage gain and the likelihood of graduation delay. The lower panel shows that the mean difference between the wage in the secondary employment spell and the employment at training end was higher for enrollees delayed than not delayed. The Wilcoxon rank sum test of the equality in distributions compares the rankings of the differences by delay category to a ranking generated by chance. This difference is significant at the .002 level. Thus the covariation of graduation delay and the wages in secondary employment spells appears consistent with a graduation strategy that maximizes the wage and earnings performance outcomes.

5 Analysis of Gaming

The above results illustrate that some training agencies time the graduation date to increase their performance outcomes and to maximize their intertemporal award streams. This section presents a simple two-task model to estimate the efficiency impact of these timing activities. To motivate this model and the estimation procedure, we start with some informal evidence that the timing activities may indeed be costly.

Timing activities may lower the organizational objective by consuming substantial resources. This is true both for graduation timing and for year-end timing. Consider first the optimal timing of the graduation date to improve the employment performance outcomes. Optimal graduation requires an up-to-date assessment of the employment state of program participants. Because of the transient nature of many employment spells and the shortness of the ninety-day window, case workers must aggressively monitor the enrollees' employment status. Such an intensive monitoring may divert resources (principally the caseworkers' time) from other productive activities, thus diminishing training's human capital impacts. In a phone survey of training agencies conducted by the authors, training managers offered details of such costs.[20] Training administrators related that they

[20]We conducted a telephone survey of the staff of thirteen training agencies; eleven of these training

find it difficult and costly to keep track of participants during the course of the training period and after the training period is over. Case workers closely monitor participants by phone, and when that is unsuccessful, they pursue them by mail, through their relatives, and even by physically staking out their last known residences.

From a welfare standpoint, training agencies clearly waste resources if these efforts are merely to obtain credit for employment that participants find on their own. On the other hand, training agencies also offer 'quick fix' services such as job referrals and placement to those individuals who were not employed after the end of training. Actually, these last-minute performance-driven services may be an efficient way to increase human capital. In the end, the issue of whether timing activities are evidence of gaming is an empirical one that can be resolved only by estimating their human capital impact.

The same argument also applies to the year-end timing activities that take place in June. In June, the burdens of graduation timing may multiply because the agency must calculate how different graduation strategies affect its current year performance relative to each of the standards it faces, as well as forecast how different graduation strategies influence its award prospects in the subsequent year. In developing a winning strategy, for example, a training agency may have to audit its inventory of enrollees, and to assess the short and long-term employment prospects of the idle portion of the case load.

To test for gaming, we will adopt the multi-tasking framework and assume that caseworkers can allocate their time between two types of activities (or tasks): *human capital activities* and *timing activities*. Under the gaming hypothesis, timing activities consume caseworkers' time and attention but do not increase human capital. Because these two activities compete with one another, we conjecture that average earnings among enrollees in any given period are lower when the caseworkers spend more time on timing activities. We ask the question: Do those enrollees whose training takes place when training agencies are more involved in timing activities receive less effective training?

agencies were from the group of sixteen training agencies that appeared in the NJS. We spoke to supervisors, managers, and training agency directors. We circulated a written questionnaire to these agencies then received their answers over the phone. The objective of the survey was to develop background information on how case workers assign enrollees to training activities, monitor enrollees while they were in training, and decide when to graduate enrollees.

Our empirical strategy will be to estimate whether human capital impacts are inversely related to timing activities. Following the job training evaluation literature, we will measure job training's human capital impact by its impact on the enrollee's earnings. In the next two subsections, we test whether earnings gains from job training depend on year-end timing and graduation timing.

Efficiency Impact of Year-End Timing. Because year-end timing only occurs in the month of June, it is more likely to lower the human capital impact of those enrollees whose training spells overlap with June. We test whether year-end timing is inefficient by asking whether training that takes place in June is less effective than in other months. For each enrollee, we construct $JUNE$, a variable equal to the fraction of her training spell spent in the month of June (e.g., for a person who started training on April 1, 1987 and finished on June 30, 1987, $JUNE$ is one third). $JUNE$ measures the extent to which enrollees may have suffered from year-end timing. Assume that earnings after training are generated by,

$$Y = \beta_1 X + \beta_2 JUNE + \epsilon \tag{7}$$

where Y is the enrollee's earnings over the 18 months following random assignment and X is a vector of control variables (including enrollee and labor market characteristics, training agency and training type dummies, and training duration). We estimate this equation using least squares and all adult data in all training agencies.[21] Under the gaming hypothesis, we conjecture that those enrollees trained in June have lower earning impacts, that is $\beta_2 < 0$.

The results of the estimation of (7) are reported in Table 4, Model I. We find that the year-end timing effect is both negative (-3959.56) and significant (the p value from the test of the estimate's significance is 0.023). The implication of this is illustrated by the example of an enrollee whose training lasts five months (roughly the average length of a training spell) and ends on June 15. The coefficient estimate implies that if the enrollee were to begin her training 15 days earlier, thus ending her training not in the middle of

[21] We use only the enrollees who report receiving training over the enrollment spell to estimate (7). See footnote 11.

June but on May 31, her 18 month earnings gain would be \$396 greater. To put this in perspective note that the point estimate from the National JTPA Study of the 18 month earnings impact ranges from \$543 to \$898.[22] That is, by beginning her training earlier so that it ended before June, this enrollee would likely enjoy a 44 to 73 percent increase in her 18 month earnings gain from training.

The negative June effect shows that those enrollees who are trained in June do not gain as much from JTPA training. This is consistent with our hypothesis that year-end timing diverts resources from human capital activities. This interpretation of the coefficient on the $JUNE$ variable, however, is subject to some caveats. To start, this finding may represent, partly or wholly, a seasonal effect. In June, JTPA enrollees may be competing with low-skilled high school and college students on summer recess. We have tested whether the effect we are attributing to June timing is instead a summer effect. We computed for each enrollee the measures $JULY$ and $AUGUST$, respectively the fractions of the training spell that occurred in the months of July and August. We added $JULY$ and $AUGUST$ to our regression. The coefficient estimates for $JULY$ and $AUGUST$ were statistically insignificant at conventional significance thresholds, and our estimate for the $JUNE$ coefficient remained essentially unchanged.[23]

Of greater concern is that the June effect may represent the burden of fiscal year-end duties, duties that do not add any new costs specific to the use of incentives. A case closing may require additional time to execute a review, file paperwork, or to complete record keeping. If costs of closing a case are incurred at the point in time the case is closed, then a disproportional fraction of these costs will be counted in June. To separately identify the incentive from the fiscal year-end effect, we considered exploiting the fact that June earnings gains would likely depend on how the agency performs at the end of the year under the gaming hypothesis but not under the fiscal year-end hypothesis. The argument

[22] These numbers are based on the figures in Exhibit S.1 of Bloom, et al. (1993). The 18 month earnings gain is the difference in mean earnings in the 18 months following the date of random assignment between persons in the control and treatment groups. See Bloom, et al. for details.

[23] The coefficient estimates (t statistics) for the variables $JULY$ and $AUGUST$ were -930.09 (-.49) and 2941.34 (1.45), respectively. We could not reject the hypothesis that the coefficients were jointly insignificant (p value was .348).

is that under the gaming hypothesis, the agency would choose how much to invest in June gaming activities by solving an inter-temporal award maximization problem that depends on its June performance and on the its June inventory of enrollees. Unfortunately, we have very poor information on the latter variable and little can be said in general about how June earning impacts should depend on the former variable alone.

Graduation Timing. Next, we test the impact of graduation timing on earning impacts. The central idea is that earnings impacts should be greater when fewer enrollees compete for the caseworker's attention. We again exploit the multi-tasking idea that time spent on graduation timing activities is time taken from human capital activities. We presume that the burden of this kind of timing increases with the number of enrollees whose graduation is being delayed. Under the gaming hypothesis we should find that, holding constant the rate at which new enrollees are assigned to a case worker, the greater the number of enrollees whose graduation is delaying, the smaller the earnings gain enjoyed by that case worker's average enrollee.

The data from the experiment are not sufficiently complete to compute accurate measures of the number of enrollees who are idle or in active training for each training day.[24] Instead, we construct a variable that proxies for the intensity of timing activities at the training agency level. The variable $DELAY$ is computed for each training agency and measures how much the agency invests in graduation timing activities relative to training activities. To compute this variable, we measure for each enrollee an index of delay equal to the number of days graduation is delayed past her enrollment spell divided by the number of days spent in active training. The variable $DELAY$ is equal to the median of that index over all enrollees in the training agency.[25] This index for the 16 training agencies ranges from 0 to .45. Agencies with an index of 0, for example, graduate at least half their enrollees right after their training ends. We assume that earnings after training are generated by the same equation as in (7) but now add the $DELAY$ variable.

[24]Unfortunately, our data contain only persons who were enrolled during the experimental window, which excludes a significant fraction of enrollees for most calendar days.

[25]We used median instead of mean because the median is less sensitive to outliers.

Under the gaming hypothesis, we conjecture that the coefficient on the $DELAY$ variable is negative.

The estimates in Table 4 confirm this hypothesis. Models II and III include on their right hand sides the measure $DELAY$, first alone and then together with $JUNE$. In both cases the coefficient estimates are negative and significant. As in Model I, the coefficient estimate for $JUNE$ in model III is negative and significant. One may argue that those training agencies that delay graduation longer are more likely to operate in depressed labor markets. This could generate a finding that earning gains are lower for training agencies that delay graduation. Models II and III therefore include labor market controls. Even after controlling for labor market environments, we find that year-end timing activities and graduation timing activities are negatively correlated with earning gains. This evidence is consistent with the hypothesis that these activities are gaming strategies, because they inflate agency performance while lowering organizational performance.

6 Summary

This paper studies timing responses to incentives. We find that JTPA training agencies time the reporting of the trainee's performance outcomes to maximize their incentive awards. We show that training agencies report good training outcomes promptly but wait on bad ones in the hope that they improve. In those cases where bad outcomes are unlikely to improve, training agencies report these outcomes only in good years where they do not risk losing their awards. In bad years, they postpone reporting bad performances until the next year to secure the minimum level of performance required to qualify for an award.

We formally test for the impact of these timing responses on the efficiency of the organization. We find that the timing responses in JTPA lower the effectiveness of job training, as measured by the impact of training on enrollee earnings. This efficiency test suggests that this timing behavior is evidence that training agencies game the incentive system. This evidence of gaming demonstrates that there are costs to using explicit performance incentives.

We conclude with a note on our perspective as to where our work fits within the incentive literature. This paper presents evidence that is consistent with the hypothesis that performance incentives in organizations lead to costly distortions in agent behavior. One could argue, however, that the costs we identify are specific to a poorly designed incentive system, and that these costs could be easily reduced or even eliminated under alternative designs. Whether these costs can easily be eliminated, however, is entirely an empirical issue that cannot be addressed until one observes and studies the responses to these alternative designs. Only by studying the responses to alternative designs can we tackle the fundamental issue of whether incentives are second-best efficient. Unfortunately, our data do not allow us to address this second-best question. Keeping this broader perspective in mind, this paper constitutes the first step toward a more ambitious efficiency analysis of the use of incentives in organizations.

References

[1] Anderson, K., Burkhauser, R. V., and Raymond, J. E. (1993). The Effect of Creaming on Placement Rates under the Job Training Partnership Act. *Industrial and Labor Relations Review*, 46(4):613–24.

[2] Asch, B. J. (1990). Do Incentives Matter? The Case of Navy Recruiters. *Industrial and Labor Relations Review*, 43:89S–106S.

[3] Baker, G. P. (1992). Incentive contracts and performance measurement. *Journal of Political Economy*, 100(3):598–614.

[4] Barnow, B. (1992). The Effects of Performance Standards on State and Local Programs. In Manski, C. and Garfinkel, I., editors, *Evaluating Welfare and Training Programs*. Harvard University Press.

[5] Brown, K., Harlow, W., and Starks, L. (1996). Of tournaments and temptation: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance*, pages 85–110.

[6] Chevalier, J. and Ellison, G. (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 105(6), pages 1167–1200.

[7] Courty, P. and Marschke, G. (1997). Measuring Government Performance: Lessons from a Federal Job-Training Program. *American Economic Review*, 87(2):383–388.

[8] Courty, P. and Marschke, G. (2000). The JTPA Incentive System: Program Years 1987-1989. In *Performance Standards in a Government Bureaucracy: Analytic Essays on the JTPA Performance Standards System*, (J. Heckman, ed.), Upjohn Monograph Series (forthcoming).

[9] Cragg, M. (1997). Performance Incentives in the Public Sector: Evidence from the Job Training Partnership Act. *Journal of Law, Economics and Organization*, 13(1):147–168.

[10] Dickert-Conlin, S. and A. Chandra. (1999) Taxes and the Timing of Births. *The Journal of Political Economy*. 107, (1): 161-177.

[11] Dickinson, K., West, R., Kogan, D., Drury, D., Franks, M., Schlichtmann, L., , and Vencill, M. (September, 1988). Evaluation of the Effects of JTPA Performance Standards on Clients, Services, and Costs. Manuscript, National Commission for Employment Policy, Research Report No 88-16.

[12] Doolittle, F. and Traeger, L. (1990). *Implementing the National JTPA Study*. Manpower Demonstration Research Corporation, New York.

[13] Drago, R. and Garvey, G. (1998). Incentives for Helping on the Job: Theory and Evidence. *Journal of Labor Economics*, 16(1):1–25.

[14] Gibbons, R. (l997). Incentives and Careers in Organizations. In, Kreps and Wallis editors, *Advances in economics and econometrics: Theory and applications: Seventh World Congress*. Volume 2. Econometric Society Monographs, no. 27, pages 1-37. Cambridge University Press.

[15] Healy, P. (1985). The Effect of Bonus Schemes on Accounting Decisions. *Journal of Accounting and Economics*, 7:85–107.

[16] Goolsbee, A. (2000). What Happens When You Tax the Rich? Evidence from Executive Compensation, *Journal of Political Economy*, forthcoming, April.

[17] Heckman, J. J. and Smith, J. A. (July, 1995). The Performance of Performance Standards: The Effects of JTPA Performance Standards on Efficiency, Equity and Participant Outcomes. University of Western Ontario, Papers in Political Economy: 73.

[18] Heckman, J. J., Smith, J. A., and Taber, C. (1996). What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program. NBER Working Paper: 5535.

[19] Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7(0):24–52.

[20] Jensen, M. C. and Meckling, W. H. (1992). Specific and General Knowledge and Organizational Structure. In Werin, L. and Wijkander, H., editors, *Contract Economics*, chapter 9, pages 252–71. Cambridge, Mass. and Oxford: Blackwell.

[21] Johnston, J. W. (1987). *The Job Training Partnership Act: A Report by the National Commission for Employment Policy*.

[22] Kerr, S. (1975). On the folly of rewarding a, while hoping for b. *Academy of Management Journal*, 18:769–83.

[23] Lawler, E. E. (1982). *Pay and Organizational Strategies and Pay Systems*. Jossey-Bass, San Francisco.

[24] Marschke, G. (2000). Performance Incentives and Bureaucratic Behavior: Evidence from a Federal Bureaucracy Manuscript, State University of New York, Albany.

[25] Murphy, K.J. (1998) Performance Standards in Incentive Contracts. Manuscript, Marschall School of Business, University of California.

[26] Orr, L. L., Bloom, H. S., Bell, S. H., Lin, W., Cave, G., and Doolittle, F. (March 1994). *The National JTPA Study: Impacts, Benefits, and Costs of Title II-A*. Abt Associates Inc.

[27] Oyer, P. (1998). The Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality *Quarterly Journal of Economics*; 113(1), pages 149–85.

[28] Prendergast, C. (March, 1999). The provision of incentives in firms. *Journal of Economic Literature*, 37(1).

FIGURE 1
Graduation Delay By Employment Status at Training End

Note: Data from the National JTPA Study. 2350 usable observations represented by graph. Plot was smoothed using a cubic spline.

FIGURE 2
Graduation Delay By Employment Status at Graduation

Note: Data from the National JTPA Study. 2350 usable observations represented by graph. Plot was smoothed using a cubic spline.

FIGURE 3
The Graduation Decision

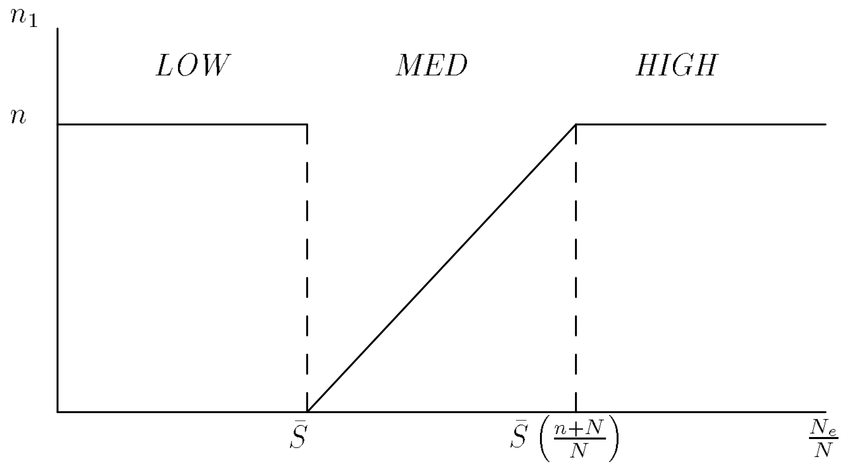
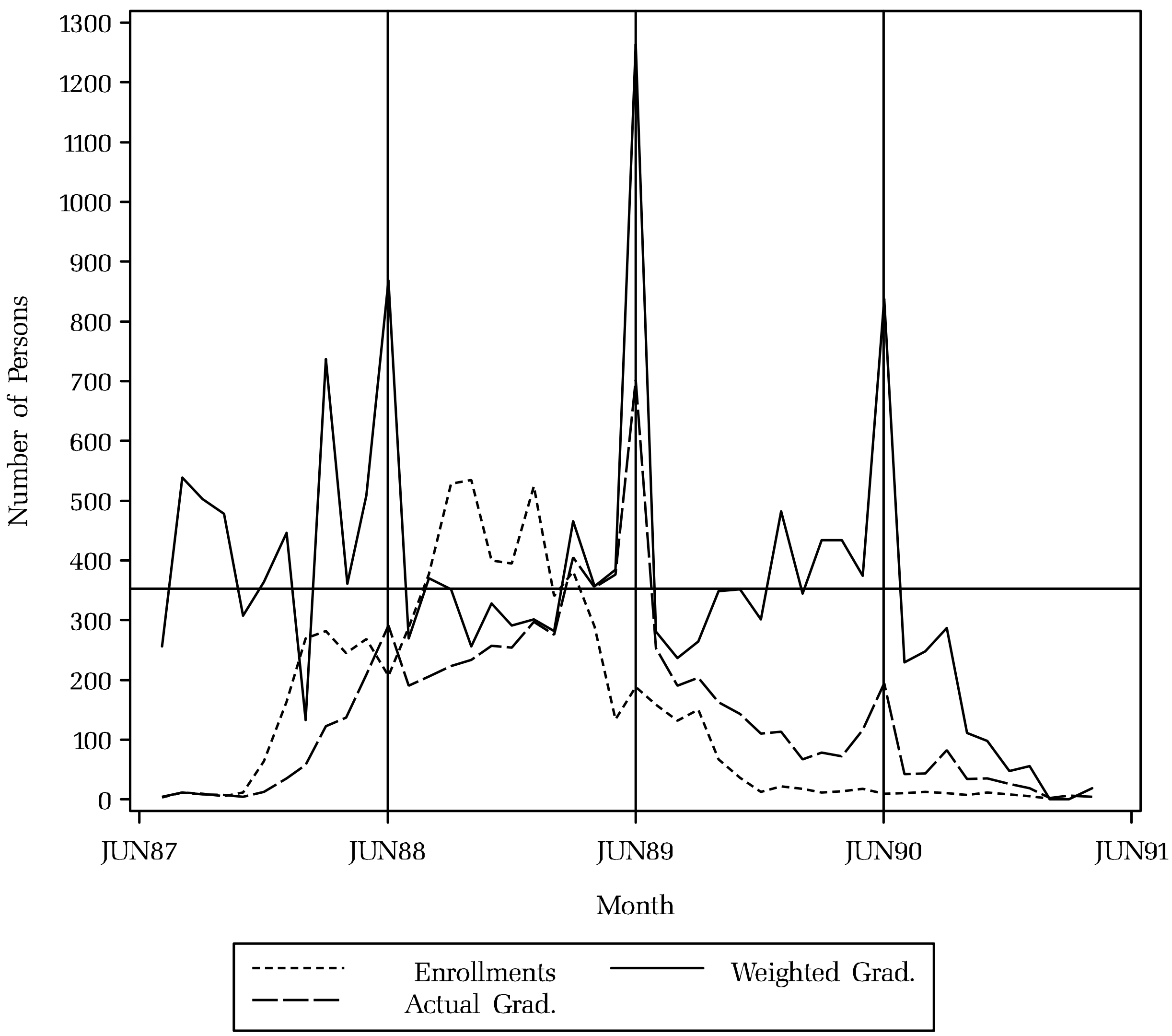


FIGURE 4
Monthly Enrollment and Graduation Counts



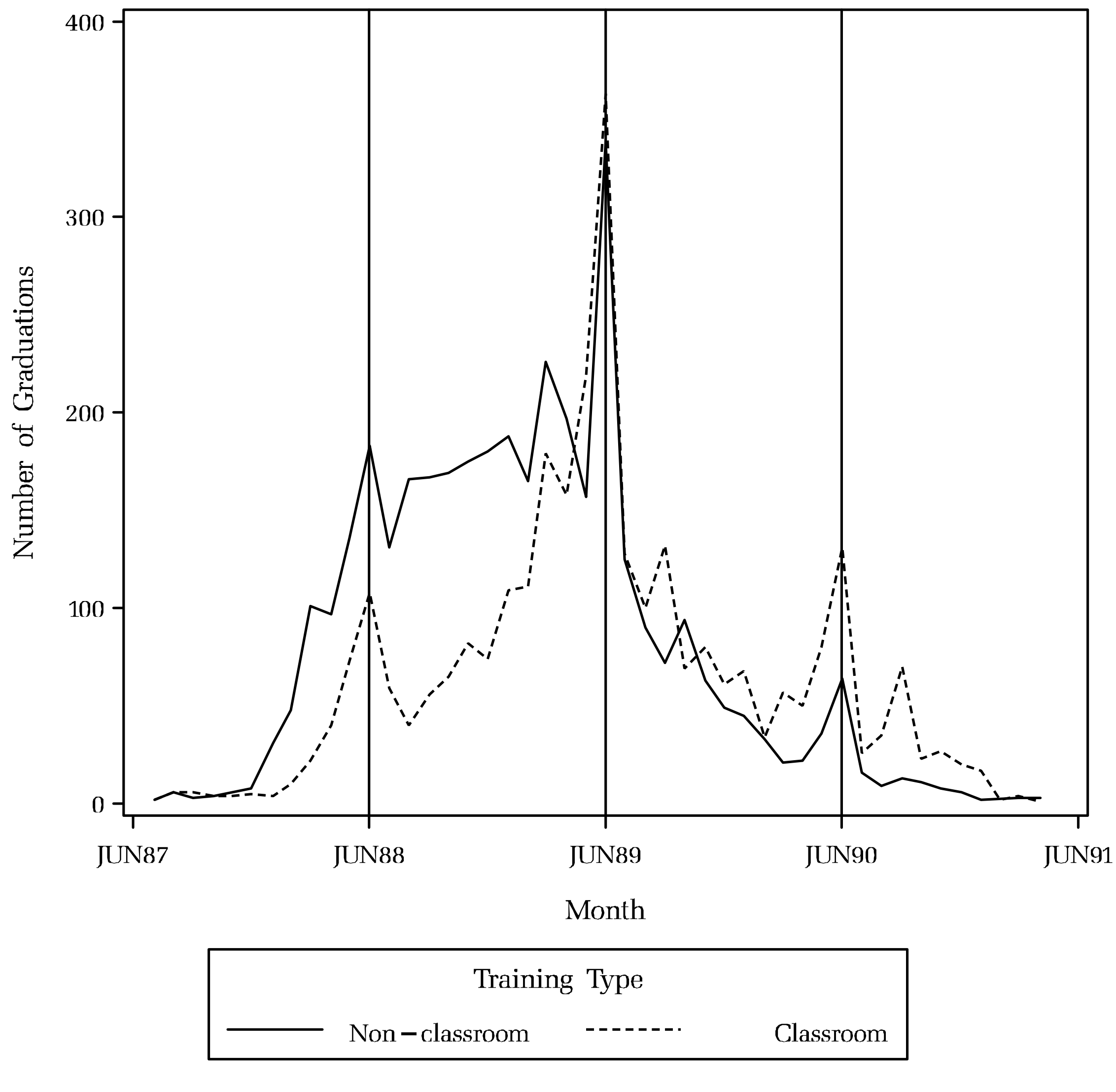
Notes: 1. Data from the National JTPA Study. 6633 adult graduates represented by graph: 903, 3775, 1704, and 296 graduates from program years 1987, 1988, 1989, and 1990. 6633 adult enrollees represented by graph: 1541, 4379, 651, and 62 enrollees from program years 1987, 1988, 1989, and 1990, respectively.

2. Enrollees in program years 1986, 1987 and 1989 reweighted to replicate the enrollment pattern of 1988.

3. Horizontal line represents mean monthly graduation using weighted sample.

4. Note that June 1988 ends program year 1987, June 1989 ends program year 1988, and June 1990 ends program year 1989.

FIGURE 5
Monthly Enrollment and Graduation Counts By Training Type



Notes: 1. Data from the National JTPA Study. Graph produced from 6677 enrollees who graduated between July 1, 1987 and May 1, 1991.

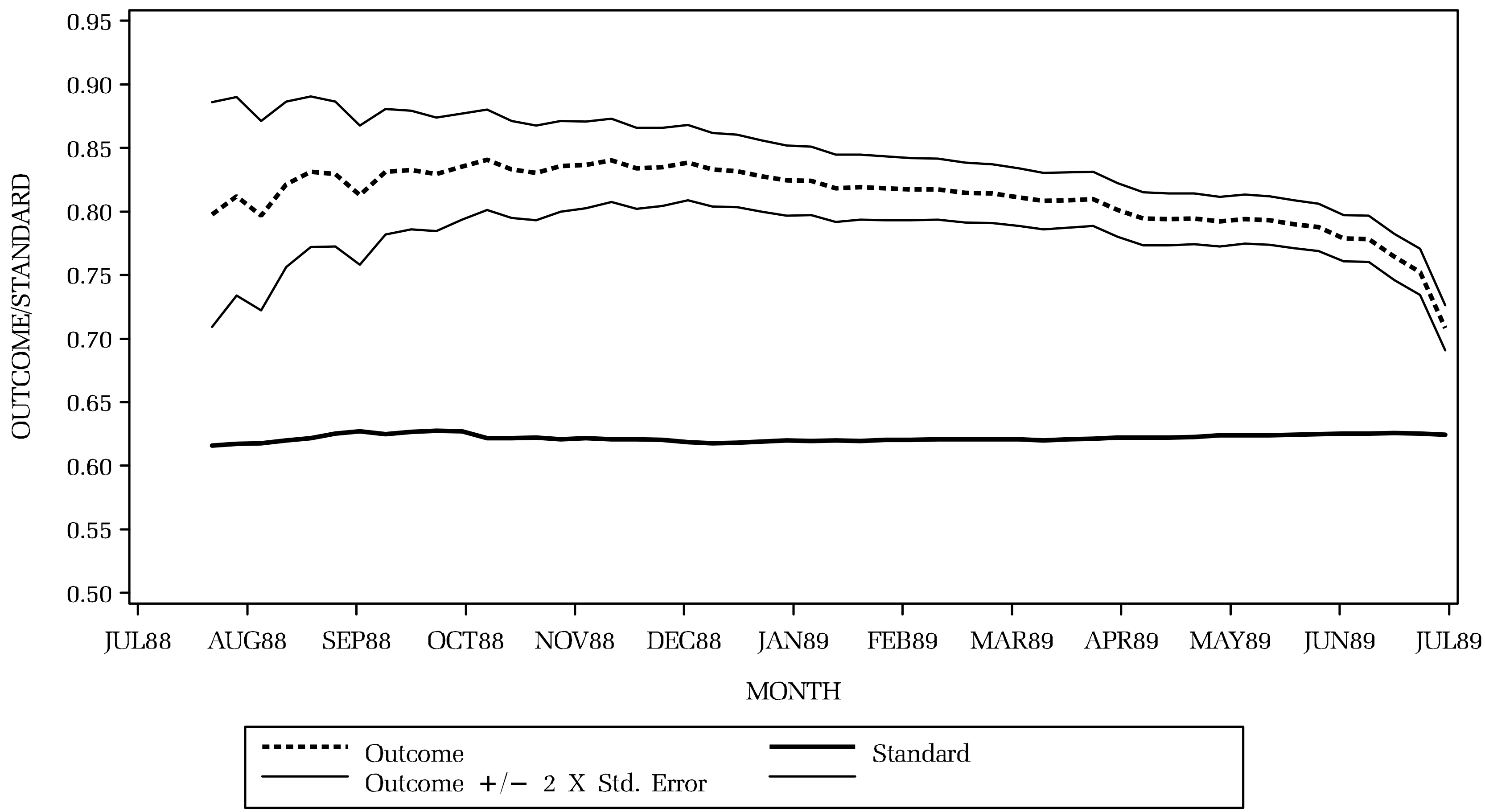
2. Enrollees in program years 1986, 1987, and 1989 reweightedto replicate the enrollment pattern of program year 1988.

3. Training type is classified according to training assignment.

4. Note that June 1988 ends program year 1987, June 1989 ends program year 1988, and June 1990 ends program year 1990.

FIGURE 6

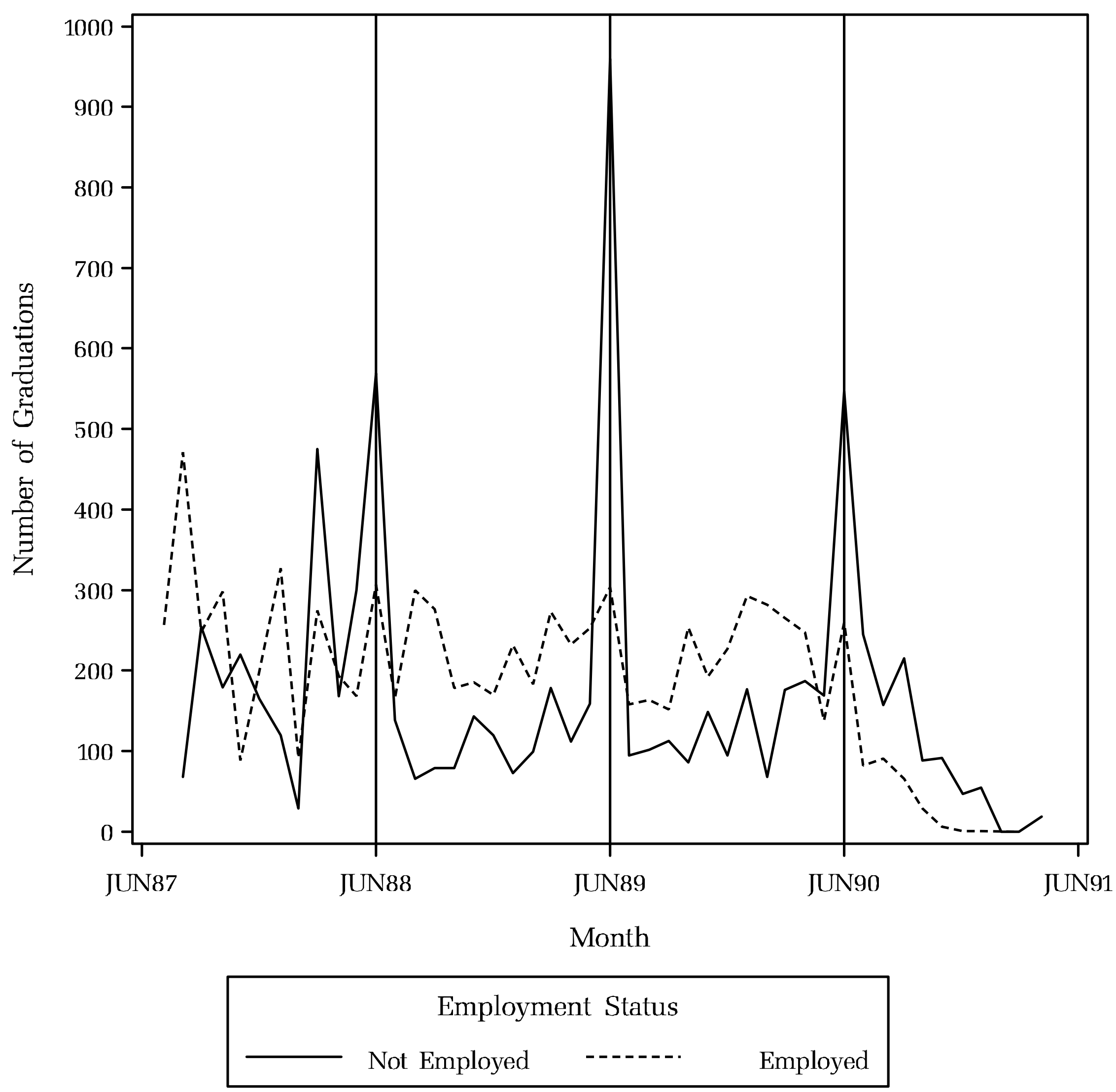
Employment Rate at Graduation in Program Year 1988 for MED/HIGH States



Note: 1. Data from the National JTPA Study. 2597 observations represented by graph.

2. All training centers whose estimated employment exceeded the standard on May 31, 1989 included.

FIGURE 7
Monthly Enrollment and Graduation Counts By Employment Status



Notes: 1. Data from the National JTPA Study. Graph produced from 6677 enrollees who graduated between July 1, 1987 and May 1, 1991.

2. Enrollees in program years 1986, 1987 and 1989 reweighted to replicate the enrollment pattern of program year 1988.

3. Note that June 1988 ends program year 1987, June 1989 ends program year 1988, and June 1990 ends program year 1990.

TABLE 1
Least Squares Estimation of Model of Graduation Decision
(t statistics in parentheses)

Dependent variable = $(\tilde{g}-\tilde{f})$

A. Coefficient Estimates

	Variable	Model I	Model II	Model III	Model IV
[0.]	$(1-u)$	34.378 (10.876)			
[1.]	$(1-u)\times(\tilde{t}-\tilde{s})$		.986 (10.061)	.798 (10.061)	.975 (4.827)
[2.]	u	101.771 (14.999)	101.771 (14.999)		
[3.]	$u\times(1-j)$			113.133 (13.837)	89.904 (6.171)
[4.]	$u\times j$			79.384 (6.629)	
[5.]	$u\times j\times LOW$				38.000 (3.001)
[6.]	$u\times j\times MED$				117.211 (33.587
[7.]	$u\times j\times HIGH$				24.791 (.575)
Obs.		2327	2327	2327	437
R^2		.1579	.1503	.1534	.1363

B. Definition of Variables:

$\tilde{g}$ = graduation date

$\tilde{f}$ = training end date

$\tilde{e}$ = first day of employment following training end date

u = 1, if enrollee was not employed anytime during the 90 days following training end, 0 otherwise

j = 1, if enrollee's training ends in months of April, May or June, 0 otherwise

LOW = 1, if training agency performance just before year-end is below performance standard, 0 otherwise (see text for precise definition of LOW, MED, and $HIGH$)

MED = 1, if training agency performance just before year-end is above but close to performance standard, 0 otherwise

$HIGH$ = 1, if training agency performance just before year-end is well above performance standard, 0 otherwise

(Continued)

TABLE 1 (Continued)

C. Hypothesis Tests[†]				
H_0	*Prob* > F			
	Model I	Model II	Model III	Model IV
1. Among idle enrollees, the duration of graduation delay is the same whether or not the enrollee experiences employment during the ninety days following training end.				
$\beta_0=\beta_2$	.0000			
2. Idle enrollees who eventually experience employment in the 90 days following training are graduated at the onset of employment.				
$\beta_1=1$		.8880	(same as II)	.9000
3. Idle enrollees who never experience employment in the 90 days following training are graduated on the 90th day following training end.				
$\beta_2=90$		.0829		
$\beta_3=90$			.0047	.9947
4. Among idle enrollees who never experience employment, graduation delay is longer when training ends in the first 9 months as compared to the last 3 months of the program year.				
$\beta_3=\beta_4$			.0200	
5. Among idle enrollees who never experience employment, whose training ends in the last 3 months of the year, the duration of graduation delay is the same whether the performance just prior to year-end is much less than (LOW=1), much greater than (HIGH=1), or greater than but close to the standard (MED=1).				
$\beta_5=\beta_6$				.0243
$\beta_7=\beta_6$				.0295
$\beta_5=\beta_7$				.7107

Notes:

† β_x is the coefficient estimate on variable no. x, where x corresponds to the number in brackets in Panel A.

1. P values are for two-tail tests. Divide by two to obtain p values for one-tail test.

2. t statistics based on robust standard errors.

TABLE 2

Performance Outcomes of Adult JTPA Clients Graduated in June and Non-June Months
All Training Centers, 1987-1989

	Non-June	June	Wilcoxon Test Prob > |Z|
Employment Rate at Graduation[1]	0.648	0.450	.0001
	(0.006)	(0.014)	
	5169	1190	
Employment Rate at Follow-up[2]	0.534	0.449	.0001
	(0.007)	(0.014)	
	5169	1190	
Average Follow-up Weekly Earnings[3]	226.584	209.015	.0005
	(2.720)	(5.255)	
	2563	509	
Average Wage at Graduation[4]	5.780	5.669	.1892
	(0.061)	(0.120)	
	2578	501	
Average Weeks Worked at Follow-up[5]	8.321	7.540	.0004
	(0.089)	(0.202)	
	4337	886	

Notes:

1. Standard errors of the mean in parentheses. Observation counts are below standard errors.
2. Calculated from adult (over 22 years of age) experimental enrollment spells.
3. For all measures except the employment rate at graduation, employment information based on enrollee reported data. Graduation dates based on training center reported data.
4. Wilcoxon test is rank sum test of equivalence of distributions.

[1]Fraction of JTPA graduates reported employed at graduation as reported by job-training center.

[2]Fraction of JTPA graduates (self) reporting at least one job spell on-going 13 weeks after the day of graduation.

[3]Average total labor market earnings in the thirteenth week after graduation.

[4]Average hourly wage in the week of graduation.

[5]Average number of weeks worked recorded in the 13 weeks following graduation.

TABLE 3

Estimation of Graduation Rule for Enrollees who

(1) Were employed at training end

(2) Experienced a secondary employment spell

		Termination after secondary spell began		
		Yes	No	Prob $>\chi^2$
Wage increase in secondary spell	Yes	61 (64%)	35 (36%)	.147
	No	18 (50%)	18 (50%)	
				Wilcoxon Rank Sum Test of Equivalence of Distributions Prob $> \lvert Z \rvert$
Difference between wage in secondary employment and wage at training end Mean ($)		2.115	.521	.002

Notes: 1. Cell counts in upper panel are followed by row fractions in parentheses.

2. Secondary spell is a spell which begins after the month training ends, but before or during the month of the end of the 90 day grace period. In the case of multiple secondary spells, an enrollee's graduation was delayed if it occured after the beginning of the first secondary spell to occur.

TABLE 4

Efficiency Impact of Graduation and Year-End Timing

Dependent variable: enrollee total earnings in 18 months following training assignment

	Coef. Estimate (t stat.)		
Variable	Model I	Model II	Model III
Agency dummies	yes	no	no
Labor market measures	no	yes	yes
JUNE	-3959.56		-3961.65
	(-2.532)		(-2.527)
DELAY		-6196.18	-6136.61
		(-3.094)	(-3.062)
Obs.	1284	1281	1281
R^2	.22	.20	.21

Notes:

1. All regressions include race; age; sex; an indicator of whether the enrollee has a child aged 6 or younger living at home; marital status; indicators of recipiency of AFDC, food stamps, and general assistance; earnings and employment histories in the 12 months prior to training; education level; type of training; and (self-reported) duration of training.

2. Labor market measures include local measures of the average wage rate, the relative employment size of the service sector, and the unemployment rate.

3. JUNE is defined as the fraction of the (self-reported) training spell spent in June.

4. DELAY is an index of delay activities at a training agency. DELAY is equal to the number of days graduation is delayed past the end of training, divided by the number of days of training, measured for the median enrollee in the experimental sample at the training agency.

5. t statistics are based on robust standard errors.

TABLE 1.A

National JTPA Performance Measures in Effect in Years 1987-1989

Performance Measure	Definition
Employment Rate at Graduation	Fraction of graduates employed at graduation
Welfare Employment Rate at Graduation	Fraction of graduates receiving welfare at date of application who were employed at graduation
Average Wage at Graduation	Average wage at graduation for graduates who were employed at graduation
Cost per Employment	Training center's year's expenditures on adults divided by the number of adults employed at graduation
Employment Rate at Follow-up	Fraction of graduates who were employed at 13 weeks after graduation
Welfare Employment Rate at Follow-up	Fraction of graduates receiving welfare at date of application who were employed at 13 weeks after graduation
Average Weekly Earnings at Follow-up	Average weekly wage of graduates who were employed 13 weeks after graduation
Average Weeks Worked by Follow-up	Average number of weeks worked by graduates in 13 weeks following graduation

Notes:

1. The date of graduation is the date the enrollee officially exits training. A graduate is an enrollee after he has officially exited training.
2. All measures are calculated over the year's *graduate* population. Therefore, the average follow-up weekly earnings for 1987 were calculated using earnings at follow-up for the graduates who graduated in 1987, even if their follow-up period extended into 1988. Likewise, persons who graduated in 1986 were not included in the 1987 measure, even if their follow-up period extended into 1987.
3. A positive graduation is entering un-subsidized employment, attaining youth employment "competencies" (through course-work, training and/or tests in work maturity, basic education, or job-specific skills), entering non-JTPA training, returning to school full-time, or completing a major level of education.