How Do Training Programs Assign Participants to Training? Characterizing the Optimal Assignment Rules of Government Agencies for Welfare-to-Work Programs in California*

Oscar A. Mitnik
omitnik@miami.edu
University of Miami

May, 2004

Abstract

Manpower training programs seek to reduce unemployment and poverty by increasing the work-related skills, human capital and employability of the poor and disadvantaged. A great deal of attention has been paid in the literature to estimating the impacts, or returns, to these programs. Much less attention has been devoted to how training agencies assign participants to training programs, what their objectives are in doing so and how these allocation decisions vary with agency resources, the initial skill levels of participants and the prevailing labor market conditions. This paper helps to fill this void by modeling the training assignment problem faced by welfare agencies, deriving empirical implications regarding aggregate training policies and testing these implications using data from Welfare-to-Work training programs run by California counties during the 1990s. I find that county welfare agencies do not follow a simple earnings-maximization model in their training assignment decisions. The results show that, as suggested by political economy models, political variables have a strong effect on training policies, specifically with respect to human capital development training.

Keywords: Public Organizations Decision Rules; Training Programs; Welfare Programs

JEL Codes: C44, D73, I38, J24

*I am grateful to V. Joseph Hotz for his tireless encouragement and support over the last several years, and to Guido Imbens for his invaluable guidance. I have benefitted from long discussions with Carlos Ponce and Eduardo Fajnzylber on the issues in this paper. I also would like to thank Sandra Black, Moshe Buchinsky, Jeffrey Grogger, Wes Hartmann, Kathleen McGarry, Enrico Moretti, Jean-Laurent Rosenthal, Duncan Thomas, and participants in the Applied Microeconomics Workshop at UCLA for their comments and suggestions. All remaining errors are the author's responsibility. Financial support from the UC Institute for Labor and Employment is gratefully acknowledged.

Part of the data in this study was used with the permission of the California Department of Social Services. The opinions and conclusions expressed herein are solely those of the author and should not be considered as representing the policy of any agency of the California State Government.

1 Introduction

Manpower training programs seek to reduce unemployment and poverty by increasing the work-related skills, human capital and employability of the poor and disadvantaged. A great deal of attention has been paid in the literature to estimating the impacts, or returns, to these programs. The issue confronting these evaluations, though, is how to deal with the bias that results from the fact that who gets trained and what type of training they receive is determined in a non-random way.[1] The processes, or decision rules, governing these choices are the domain of program administrators and governmental agencies. Unfortunately, to date, we understand very little about these decision rules and the objectives of these agencies/administrators.[2,3]

This paper helps to fill this void in our knowledge using data for Welfare-to-Work (WTW) training programs run by counties in California during the 1990s. In 1996, the U.S. welfare system underwent a major reform with the passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA). One of the key features of the new system was to require most participants to enroll in one or more training or work-related activities, under WTW programs, as a condition of receipt of this temporary assistance.[4,5] Within these WTW programs, training is typically classified into two types: labor force attachment (LFA) programs that provide recipients with training and assistance in finding jobs, and human capital development (HCD) programs that seek to develop basic and work-oriented skills. In this paper, I attempt to understand the decisions made by the program administrators; in particular,

[1]See Heckman and Robb (1985) and Heckman, LaLonde and Smith (1999), amongst other articles, for a discussion of the issue of selection bias. (the latter paper provides an extensive survey of methods and evidence on the evaluation of training programs).

[2]However, the analysis of the determinants of government policies has been an issue of interest for a long time in economics. The areas of interest have been diverse. Classic examples are the study by McFadden (1975, 1976) on decision rules underlying freeway route selection, and the book by Wilson (1989) on several aspects of the functioning of government agencies. Other areas of interest have been, for example, the analysis of monetary policy rules by central bankers (see the collection of articles in Taylor, 1999) and the fiscal consequences of political regimes (see for example Persson, Roland and Tabellini, 2000; Persson and Tabellini, 2002; and Besley and Case, 2003).

[3]Special attention has received recently the modeling and empirical estimation of the effects of incentives and performance standards on the decisions of public sector organizations (see Dixit, 2002, for an overview of the issues surrounding incentives in organizations in the public sector), particularly related to training and education policy. Regarding education policies, for example, see Eberts, Hollenbeck and Stone (2002), Hoxby (2002), and Koretz (2002). Regarding training policies, see Courty and Marschke (1997), Dehejia (forthcoming), Heckman, Smith and Taber (1996), Heckman, Heinrich, and Smith (1997 and 2002), and Pepper (2002 and forthcoming).

[4]Note that having a requirement for welfare recipients to participate in training or work-related activities was not new to PRWORA. Since the the 1960s, the U.S. welfare system, under the old Aid to Families with Dependent Children (AFDC) program, included requirements for participation in employment and training programs by non-exempted welfare recipients. But these provisions were not consistently enforced. The current federal legislation contains stronger mandates and incentives to states to comply with these provisions, especially the participation in work-related activities.

[5]One interpretation of the work and training requirements is that they are just a devise to identify the people really needy from the people that want to enjoy leisure (see Besley and Coate, 1992). Although this might be one of their objectives, I will assume throughout this paper that the training programs do have added value.

who these programs/agencies train and, more importantly, which type of training they provide to particular participants. To do so, I model the decision-rules for how welfare agencies run their WTW programs.

I examine several alternative models of how government agencies assign workers to different training programs. In an initial model, I assume that welfare agencies assign workers to different types of training, including a no training option, so as to maximize the well-being —or, specifically in this case, the present value of their earnings— of program participants subject to an agency's budget constraint and the costs of the alternative types of training. This "earnings maximization" model of how training programs operate characterizes a "utilitarian" view of the objectives of government agencies and is a natural place to begin to study agency/program decision-making. This is the base model in the study by Heckman, Heinrich, and Smith (2002) of how agencies respond to the imposition of performance standards. In alternative models to the earnings-maximization model, I allow for a richer set of agency preferences, both across types of training and/or across people depending on their initial skills.

In modeling the training choices of agencies, I allow for several important sources of heterogeneity with respect to participants and the labor markets in which they operate. First, I allow for the possibility that program participants are heterogeneous with respect to their initial skill levels. Second, I also allow for agencies to differ with respect to the labor market environments for which they are training participants, and for the possibility that the nature of labor markets interact with participant skill levels to affect the marginal returns to earnings for different types of training. As a result, I allow for the possibilities that training is not beneficial for every individual (i.e. is optimal not to train some people), that any particular type of training does not benefit all individuals in the same way, and that the effectiveness of different training programs varies with the conditions prevailing in the labor market. Also, in my analysis, I explore alternative assumptions about how the training production functions, i.e., the mappings of training type to subsequent earnings, vary with initial skills and labor market conditions.[6,7]

[6] More formally, I allow for skill and labor market heterogeneity that induces heterogeneity in the marginal returns to different types of training. As such, the model shares the essential features of the Roy Model of self-selection in the labor market (Roy, 1951; see also Willis,1986 and Heckman and Honore, 1990), in which heterogeneous agents self-assign themselves to occupations according to a principle of comparative advantage. Both sources of heterogeneity, and their consequences for optimal assignments to training activities, are thought to be important in the context of the allocation of training (see, for example, Heckman and Robb, 1985, and Pepper, forthcoming).

[7] Note that I assume that training program administrators know the expected marginal returns to earnings of different types of training for each participant. Thus, I ignore an issue, namely that program administrators may lack information about the potential impacts of training for program participants before assignment. When facing less than complete information about such returns, program administrators must make inferences about the expected returns to training if they are to use it in their training assignment decisions. This issue is considered by Pepper (forthcoming), Manski, Newman and Pepper (2002), and Dehejia (forthcoming).

My primary focus is on how different objectives attributed to the program agencies affect their assignment decisions. As noted above, I first consider the earnings-maximization model of agency training assignment. Under this model, the agency will allocate participants to the alternative types of training to achieve efficient allocations that equalize marginal benefits and marginal costs of training. In the special case where the agency's budget constraint is not binding, it corresponds to an individual participant's problem where she optimally self-selects the type of training she undertakes as if she had access to perfect capital markets. In this special case, the agency problem reduces to maximizing the present value of the marginal impacts of training for each participant, implying that the role of the training program is to overcome the imperfections of private capital markets.

To the extent that the budget constraints of the welfare agency are binding, the optimal assignment of individuals to treatment will depend on what is analogous to the social marginal costs of training. Furthermore, the optimal assignments will differ from the unconstrained case in that the optimal policies of agencies will ration-out some individuals from any training and for some other individuals will substitute from more expensive training to cheaper training, until the relative returns of the treatment alternatives are equated to the relative costs. Note that the problem of how the agency assigns participants to training, subject to a common budget constraint, has close parallels to the intrahousehold allocation decisions of parents when deciding how much to invest in the education of their heterogeneous-in-skills children.[8]

Finally, I consider the possibility that welfare programs have a richer set of preferences over participants than implied by a simple earnings-maximization model. Because oversight of welfare agencies is done by elected officers, a political economy model suggests that public opinion and the beliefs of the median voter should affect the assignment to training policies.[9] As is well documented in various types of polling data, citizens often have strong views about people on welfare, the appropriateness of having a welfare system and what types of benefits and services people on welfare should receive.[10] Also, political parties and quasi-public organizations (e.g., unions and special interest groups) often maintain explicit positions, or platforms, relative to the structure of welfare programs.[11] Lastly, welfare agency administrators and case-

[8]See Becker and Tomes (1976), Behrman, Pollak and Taubman (1982) and Behrman (1997). The models of the family in these papers differ from the training problem in that parents seeking to maximize the wealth of their (dynastic) family seek to make decisions about both educational investments and financial transfers (either bequests or inter-vivo transfers), whereas in the assignment to training models I analyze agencies only consider how to allocate individuals across types of training.

[9]There is a debate in the literature on whether politicians respond to public opinion or, just to their beliefs and of their supporters. See Jacobs and Shapiro (2002) and Erikson, Mackuen and Stimson (2002) for discussion and some evidence on this issue.

[10]See Kluegel and Smith (1986), Cox and Barrett (1992), Bean and Papadakis (1998) and Weaver (2002).

[11]Various theories have been put forward in the political science literature concerning how policy decisions related to the distribution of services to targeted groups, such as those on welfare, might be affected, including

workers, may develop their own independent objectives for how their programs are run and which individuals should be served.[12] I generalize the earnings-maximization model to allow for a richer set of agency preferences, including different preferences across types of training and/or across people depending on their initial skills. I show that these preferences lead to different assignment decisions than the ones that result from simple earnings maximization.

Based on these models, I derive a number of testable empirical implications. In particular, I derive comparative-static predictions concerning how the proportions assigned to different types of training (including a no-training option) vary as one varies the distribution of initial skills of the participant pool in a particular administrative locality, the budgets for that locality, and the labor market conditions in that locality. I derive these predictions under minimal assumptions about the nature of the training production functions. Imposing more structure on these production functions at the theoretical level can yield more precise comparative static results, but they are made at the cost of having these predictions depend on maintained assumptions that cannot, themselves, be tested with the data available. Thus, it is preferable to focus the attention on a robust, but more limited, set of predictions in my empirical analyses.

As noted above, I test the implications of my alternative models using county level panel data for the WTW programs of counties in California.[13] California is a well suited application because it runs a very decentralized welfare system in which individual counties have great latitude over which training schemes they offer and how they assign welfare recipients across these programs. The focus of the empirical analysis is on how the proportions of welfare recipients that received three types of treatment—no training, labor force attachment training and human capital development training—vary as a function of the budgets they receive from the State to run their training programs, the characteristics of their welfare recipients and indicators of the local labor market conditions prevailing in the county. As it is shown below, there is a great deal of diversity across California's counties in their pool of welfare recipients and their labor market conditions over the 1990s.

One difficulty with the empirical analysis involves characterizing the unobserved distribution of initial skills of the potential trainees. Although I can not observe actual skills, I am able to use microdata to approximate them. This is accomplished by using individual level information on welfare recipients and constructing a measure of skills as the residual from an

---

theories of self-interest (i.e., individuals that are more likely to benefit from a policy support it) and political predisposition (for which party identification has been shown to be a good predictor). See Kluegel and Smith (1986) and Cox and Barrett (1992).

[12]Evidence on the effects of preferences of caseworkers is presented by Heckman, Smith and Taber (1996) regarding acceptance into the JTPA training program.

[13]During the period analyzed (1994-1999) training in California was offered to welfare recipients through the GAIN Program, which was renamed as WTW in 1998, when TANF started in California. Section 2 explains in details the characteristics of the programs.

econometric model of labor force participation and earnings. In a second step, aggregate statistics from the skills distribution, along with county level data on expenditures, local economic conditions and registration and voting patterns are used to explain the variation in training policies.

My findings imply that a simple earnings-maximization model does not characterize the observed assignment decisions of county welfare agencies. In particular, I find that political variables, such as voting and registration patterns, have a strong effect on training policies, specifically positive with respect to human capital development training. This is consistent with political economy models of assignment to training in which decision makers have preferences for either a particular type of training or a particular group of participants. These results suggest that the often assumed objective of earnings-maximization in the literature needs to be reconsidered. They also imply that government agencies exhibit complex objective functions that need to be taken into account in the design of policies.

The paper is organized in seven sections. The next section describes the WTW training programs for welfare recipients in California, and shows the evolution over time of these two types of training. The third section presents the optimization problem faced by a welfare office that tries to maximize the expected outcome from its training policies and derives the empirical implications of such a problem. It also analyzes extensions to the basic model. The fourth section presents the county level and individual data to be used. The fifth section describes the empirical strategy, which entails estimating measures of the distribution of observable and unobservable characteristics of welfare entrants in a first step and estimating regressions using the county level data on training policies in a second step. The sixth section presents the results, while the seventh section concludes.

2 The GAIN and WTW programs

Training was offered to welfare recipients in California in the 1990s through two programs. California's version of the JOBS Program[14], under the Aid for Families with Dependent Children (AFDC) Program, was the Greater Avenues for Independence (GAIN) Program, in which training was the main component. It started in 1989 and it was succeeded in 1998 by the Welfare to Work (WTW) Program, as part of the California Work Opportunity and Responsibility to Kids (CalWORKs) program, California's version of the Temporary Assistance for

[14] The Job Opportunity and Basic Skills (JOBS) Program was directed to help families on welfare avoid long-term welfare use, by providing job search assistance, education, work experience, vocational training, and other employment-related services, and required all parents (except those with small children) to participate in these work-related activities or face a reduction in the amount of assistance received (Haveman and Wolfe, 2000).

Needy Families (TANF) Program.[15] Both programs offered different types of services to welfare recipients who were mandated to participate (except parents of small children), or face financial sanctions. However, under GAIN counties faced severe funding constraints, and in some counties a big proportion of the caseload remained not served.

Under the rules of CalWORKs every adult is required to participate in the Welfare to Work Program, which implied that counties had to expand their programs to accommodate all the adult caseload.[16] The activities that the programs offered included, among others, job search and job readiness assistance, on the job training and subsidized employment, vocational education training, adult basic education, English as a second language, and classes for preparing to take the General Education Diploma (GED) exam.

The training activities have been classified in two groups: work-oriented, termed Labor Force Attachment (LFA), and education-oriented, termed Human Capital Development (HCD), types of training. Typically LFA training has a shorter duration and is less expensive to provide that HCD training (see Hamilton, Freedman et. al., 2001 for a discussion of both approaches).[17]

Counties in California were given a great degree of freedom over the design of their GAIN and WTW Programs. This caused a remarkable variation both across counties and across time in the proportion of the adult caseload that participated in any activity, and that participated in LFA and HCD types of training. The variation from 1994 through 1999 in the proportion of adults receiving any training, and the two types of training is presented in Figure 1.[18] As can be seen in the figure, there has been an increasing trend in the proportion of individuals trained in general, but more marked for LFA training. When the CalWORKs program started (January 1998), there was a larger jump in absolute terms for LFA than for HCD training, although the latter increased more in relative terms. This figure does not show the important cross county variation in these proportions, but it can be observed in the Appendix Table A.1, which presents yearly data for the 25 biggest counties in California, the ones that will be used in this study.

[15]TANF replaced AFDC after the passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) in 1996. California was the last state to implement TANF, starting in January 1998 under the CalWORKs name.

[16]However, participation in the WTW program does not mean necessarily receiving training services. There are non-training activities that count as participation in the WTW program (like using substance abuse services).

[17]The Labor Force Attachment (or Work First) type of programs try to increase the insertion of welfare recipients into the labor force by providing job search training and assistance, while the Human Capital Development type of programs are oriented to increase the trainees human capital by offering longer duration basic skills and vocational training programs.

[18]This paper covers the period January 1994-June 1999 because after June 1999 a new report system for the WTW program (WTW25) was implemented by the state of California, that was not fully functional until 2000. Because the comparability of the data from WTW25 with the data from the original report system (GAIN25) is far from clear, only the data from GAIN25 is used.

Note that the figure presents data for all the adults on welfare in a particular period of time. Unfortunately there is no data available that breaks these proportions between new entrants and non-entrants to welfare. The theoretical model and empirical analysis will be based, however, in the characteristics of new entrants to welfare, because their assignment to training is a well defined problem, while the problem in which the county has to decide the assignment to training for a welfare recipient conditional on her (potentially) having already received training before is a more complex problem. Moreover, both the GAIN and WTW programs (with more emphasis the second one) were devised such that each new entrant to welfare was assessed and assigned to a treatment in a relatively short period of time. However, under the GAIN program, counties in which the available resources did not permit the treatment of a large percentage of the caseload had a long waiting list for welfare recipients assigned to a particular training. This will be reflected in a "no training" category, which will be explicitly modeled in the next section.[19]

3 Assignment to Training Models

In this section I develop a simple earnings maximization model for the decision making process of a county that has to assign welfare recipients to different training alternatives. I characterize the solution to the problem, and derive testable implications with respect to the effects of changes in the budget constraint, in local economic conditions and in the initial skills distribution of the welfare entrants. In the last subsection I extend the basic model to allow for political economy considerations that imply that welfare programs have a richer set of preferences over participants than implied by the simple earnings-maximization model. Specifically I generalize the earnings-maximization model to consider agency preferences across types of training or across people depending on their initial skills. I show that these preferences lead to different assignment decisions than the ones that result from earnings maximization.

The hypothesis of the simple model is that the objective of the county is to maximize the expected returns from its training policies. Since the county cares about the outcomes of all individuals, trained or not, this is equivalent to maximizing expected earnings. This seems like a natural starting point, given that it reflects the view that the welfare agency's objective is to obtain the best possible outcomes in the labor market for the welfare recipients, to help them be self-sufficient. Furthermore, this objective function is the same as that assumed by Heckman, Heinrich, and Smith (2002), Dehejia (forthcoming), Manski (2000, 2001) and Pepper

[19]The "no training" treatment might mean, then, that if the individual stayed long enough in welfare, she could have received eventually some training services, even though that did not happen as a new entrant.

(2002, forthcoming).[20]

The model by Heckman, Heinrich, and Smith (2002) is similar to the one I present below, in particular, because it incorporates into the analysis a budget constraint, that the other papers do not take into account.[21] Also, their model allows the decision maker to choose the level of training effort for each individual, which in some sense is similar to the problem in this paper, in which the county has to choose among three treatment options. The contribution of the model in this paper is to show how the decisions of the county regarding the two training options can be used to empirically test implications of alternative objectives governing agencies' assignment decisions.

3.1 Setup

Assume a county needs to make training decisions regarding an entrant cohort of welfare recipients. Each individual has certain associated characteristics, denoted by a random vector X_i, which includes her demographic information, education attainment, previous employment and earnings histories, and, in some cases, evaluations (either objective or subjective) made by county officers on the individual's potential.[22] I assume that this information is aggregated by the county in a single-dimension index $\theta = \omega(X_i)$, where ω is a known weighting function. Furthermore, I assume that this index is continuous, that it takes on values in the interval $[\underline{\theta}, \overline{\theta}]$, and that it follows a county specific probability distribution function $f(\theta)$. I assume that the ability to generate earnings or to find employment is positively correlated with θ, which can be interpreted as a measure of the initial skills of the individuals. I will refer subsequently to θ simply as *skills*.

The county has three treatment options for each individual: to not provide her any training services (*treatment* N), to provide her with LFA training services (*treatment* L), or to provide her with HCD training services (*treatment* H). The county evaluates the effect of each treatment option on an individual with skills θ, by looking at the outcome variable $Y(\theta)$. This variable can represent different outcomes of interest for the county (like earnings, probability of finding a job, etc.). To make the discussion specific I assume that $Y(\theta)$ represents an earnings generating function.

Treatments N, L and H have associated, at each period, earnings generating functions $Y^N(\theta)$, $Y^L(\theta)$ and $Y^H(\theta)$ respectively. The earnings functions for treatments L and H can

[20] However, Dehejia also incorporates uncertainty regarding the distribution of outcomes, and allows for a social welfare function that exhibit inequality aversion.

[21] Pepper (forthcoming) does consider as one possible source of identification of the potential effects of WTW polices an optimization model with a budget constraint.

[22] Some elements of the vector X_i will be observed and some will be not observed. I discuss this issue in detail in section 5.

also be rewritten as

$$\begin{aligned} Y^L(\theta) &= Y^N(\theta) + \Delta^L(\theta) \\ Y^H(\theta) &= Y^N(\theta) + \Delta^H(\theta) \end{aligned}$$

where Δ^L and Δ^H represent the *training effects* of training L and H respectively, and they are assumed to be non-negative functions of θ. This assumption just implies that an individual can not have her earnings generation capacity diminished because of receiving training.[23]

However, because training takes time (and hence forgone income), the *present value* of the earnings under treatment N could be higher than the present value of the earnings under treatment L or treatment H, for some values of θ. I assume that each individual, at the moment of treatment, has a future labor life of T periods (fixed). Treatment L takes place over $\tau_L > 0$ periods, while training H takes place over $\tau_H > 0$ periods. Therefore, as HCD training takes longer than LFA training, this implies $\tau_H > \tau_L$.

Then, the expected present value of the future stream of benefits (in each period t) associated with receiving treatment i is represented by the benefit functions

$$V^i(\theta) = \sum_{t=\tau_i}^{T} \rho^t E[Y_t^i(\theta)]$$

where ρ is the discount factor, $i = N, L, H$, and $\tau_N \equiv 0$. To simplify the setup I assume that the expected earnings are constant over time, i.e. $E[Y_t^i(\theta)] = E[Y^i(\theta)]$ for all t, which implies that the benefit functions can be re-expressed as

$$V^i(\theta) = \kappa_i E[Y^i(\theta)]$$

where κ_i is a constant.[24] Without loss of generality it can be assumed that $\rho = 1$, in which case $\kappa_N = T$, $\kappa_L = T - \tau_L$, and $\kappa_H = T - \tau_H$ (and it is easy to see that $\kappa_N > \kappa_L > \kappa_H$).[25]

Because the cost of training is borne by the county, for the individuals, there is no direct cost of receiving training L or training H, other than the opportunity cost of attending the training. Also, I assume there is no cost to the county associated with the provision of treatment N (no training), i.e. $c_N \equiv 0$. Thus, the county faces direct costs per person c_L and c_H associated with the provision of treatments L and H respectively, where $c_H > c_L$, because HCD training,

[23]However, note than the differential treatment effect $\Delta^1(\theta) - \Delta^2(\theta)$ can be positive or negative.

[24]One interpretation to a constant over time $Y^i(\theta)$ is that it represents the average earnings over the period that would imply the present value of earnings $V^i(\theta)$.

[25]Note that if $\rho < 1$, then $\kappa_N = \frac{1-\rho^{T+1}}{1-\rho}$, $\kappa_L = \frac{\rho^{\tau_L}-\rho^{T+1}}{1-\rho}$, $\kappa_H = \frac{\rho^{\tau_H}-\rho^{T+1}}{1-\rho}$ and it still holds that $\kappa_N > \kappa_L > \kappa_H$.

is more expensive than LFA training.[26] Finally, the county has a fixed budget B which it has to use to cover the expenses of providing the training services.[27]

3.2 The county's problem

The county welfare agency seeks to assign individuals to the different treatments so as to maximize the expected returns to the investment. Denoting by $I_i(\theta)$, the indicator functions equal to 1 if the individual with skills θ is assigned to treatment i ($i = L, H$) and equal to 0 otherwise, the county optimally chooses these indicator functions by solving the following problem:

$$\begin{aligned}\max_{\{I_L(\theta), I_H(\theta)\}} W = &\int_{\underline{\theta}}^{\overline{\theta}} [V^L(\theta) - V^N(\theta) - c_L] I_L(\theta) dF(\theta) \\ &+ \int_{\underline{\theta}}^{\overline{\theta}} [V^H(\theta) - V^N(\theta) - c_H] I_H(\theta) dF(\theta) \qquad (P1) \\ s.t. \int_{\underline{\theta}}^{\overline{\theta}} &[I_L(\theta) c_L + I_H(\theta) c_H] dF(\theta) \leq B \\ I_L(\theta) &+ I_H(\theta) \leq 1 \quad \forall \theta.\end{aligned}$$

This problem implies that the county has to assign a treatment (choose I_L and I_H) for each individual. Although an analytical solution can be obtained, it is not a particularly interesting formulation in terms of obtaining empirical implications. To make some progress, I impose minimal conditions on the earnings generating functions Y^N, Y^L and Y^H that make the problem both analytically simpler, and more importantly, informative in terms of implications of changes in the economic environment faced by the county.

Assumption 1. The earnings generating functions $Y^i(\theta)$ $(i = N, L, H)$ are:
i) twice differentiable
ii) concave: $Y^i_{\theta\theta} <= 0$
iii) strictly increasing: $Y^i_\theta > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$
iv) bounded from below and above: $Y^i(\underline{\theta}) \geq 0$ and $Y^i(\overline{\theta}) < K$, where the constant $K < \infty$.

Assumption 1 implies that there are always positive but decreasing returns to skills for any treatment and that the problem will be well specified. Note that with Y^N concave, it is

[26] See Hamilton, Freedman et. al. (2001).

[27] In the model I abstract from the budget allocation process, by which B is available to spend in training programs. This simplification allows to focus on the determination of training policies, given a budget. However, in the empirical analysis it must be recognized that the allocation of resources to the training programs is not endogenous. This issue will be dealt with in the empirical specifications.

sufficient that the treatment effects Δ^L and Δ^H are concave, to have earnings functions Y^L and Y^H that are concave.[28]

Assumption 2. There is an ordering of the (present value of) expected marginal returns to skills for the three earnings functions Y^N, Y^L and Y^H that is unchanged for all $\theta \in [\underline{\theta}, \overline{\theta}]$. That is, for all triplets of earnings functions (Y^i, Y^j, Y^k) ordered by their marginal returns, it is satisfied that

$$\kappa_i Y_\theta^i < \kappa_j Y_\theta^j < \kappa_k Y_\theta^k \quad \forall \theta, \;\; i, j, k = N, L, H, \;\; i \neq j \neq k.$$

Assumption 2 has a clear economic interpretation: if a particular treatment is better for individuals of higher skills, then the marginal returns to skills are also higher for these individuals. This is a standard single crossing property (i.e. ensures that the benefit functions cross at most once), and is crucial to be able to simplify the decision problem of the county.[29]

Assumption 3. The cumulative distribution function of θ is differentiable and strictly increasing in θ: $F_\theta(\theta) \equiv f(\theta) > 0$, for $\theta \in [\underline{\theta}, \overline{\theta}]$.

Assumption 3 guarantees that there are no regions of θ in which the density function is not defined, and it is necessary to assure that the first order conditions of the problem are well defined.

It is straightforward to prove that under Assumptions 1 and 3 a solution always exists.[30] Moreover, Assumptions 1 and 2 imply that the interaction of the functions $V^N(\theta)$, $V^L(\theta) - c_L$ and $V^H(\theta) - c_H$ divide the support of θ, into (at most) three regions, in which all the individuals belonging to a region will receive the same treatment.

I denote the three regions over the distribution of θ as R_l (low), R_m (medium) and R_h (high), and use the region subscript to denote the treatment provided in that region (i.e. l, m and h, where each of them can assume values N, L or H). Then, the county problem reduces to choosing two critical values for θ, that define the regions' limits. Designating θ_l as the critical value that separates regions R_l and R_m, and θ_h as the critical value that separates regions R_m

[28]Although this is not necessary as long as $Y_\theta^N > \Delta_\theta^L$, $Y_\theta^N > \Delta_\theta^H$ and $|Y_{\theta\theta}^N| > |\Delta_{\theta\theta}^L|$, $|Y_{\theta\theta}^N| > |\Delta_{\theta\theta}^H|$.

[29]This condition is analog to what is known in the incentives literature as "Spence-Mirrlees" condition or "constant sign" condition (see for example Salamié, 1997, pp. 31).

[30]Assumptions 1 and 3 guarantee that the objective function W is continuous and that the budget set is compact.

and R_h, the county will choose the θ_l and θ_h that solve the following problem:

$$\max_{\{\theta_l,\theta_h\}} W = \int_{\underline{\theta}}^{\theta_l} [V^l(\theta) - c_l] dF(\theta) + \int_{\theta_l}^{\theta_h} [V^m(\theta) - c_m] dF(\theta)$$
$$+ \int_{\theta_h}^{\overline{\theta}} [V^h(\theta) - c_h] dF(\theta) - \int_{\underline{\theta}}^{\overline{\theta}} V^N(\theta) dF(\theta) \qquad (P2)$$
$$s.t.\ F(\theta_l)c_l + [F(\theta_h) - F(\theta_l)]\ c_m + [1 - F(\theta_h)]c_h \leq B,$$

where c_l, c_m, or c_h will assume the value 0 when valued at treatment N (i.e. $c_N \equiv 0$).

The formulation of $(P2)$ is more interesting than the one for $(P1)$ because it allows testable implications to be derived. Whereas in $(P1)$ the welfare agency has to make a decision tailored for each individual that is a candidate to receive treatment, in $(P2)$ the county only needs to set a decision rule based on the two critical values, and assignment to each treatment follows directly from this decision. This allows the study of the effects of changes in the environment faced by the county welfare agency by just analyzing the effects on the critical values for θ.

3.3 Solution

The first order necessary conditions for $(P2)$ (using a Lagrange multiplier λ for the budget constraint) are the following:

$$\theta_l : [V^l(\theta_l) - c_l]f(\theta_l) - [V^m(\theta_l) - c_m]f(\theta_l) + \lambda f(\theta_l)(c_m - c_l) \leq 0 \qquad (1)$$
$$\theta_h : [V^m(\theta_h) - c_m]f(\theta_h) - [V^h(\theta_h) - c_h]f(\theta_h) + \lambda f(\theta_h)(c_h - c_m) \leq 0 \qquad (2)$$
$$\lambda : B + F(\theta_l)(c_m - c_l) + F(\theta_h)(c_h - c_m) - c_h \geq 0. \qquad (3)$$

The second order conditions for a maximum are always satisfied under Assumption 2, as shown in the Appendix.

I only consider the "interior solutions" to this problem, in which the three treatments are always provided (i.e. the regions R_l, R_m and R_h are not empty).[31] Depending on which type of individuals benefits more from each type of treatment, there are six possible cases when considering "interior solutions". I will characterize these cases by the treatment received by the individuals in the regions of low, medium and high skills, in this order. Denoting them by the notation $[R_l, R_m, R_h]$, where the subscripts refer to the treatment, the cases are:

1. $[R_N, R_L, R_H] \Longrightarrow$ [No Training, Training L, Training H]

[31] Note, however,that there are configurations of the benefit functions, cost and timing of the treatments under which one or even two of the treatments might not be offered in equilibrium. These instances can be obtained as extreme cases of the interior solution ones.

2. $[R_N, R_H, R_L] \Longrightarrow$ [No Training, Training H, Training L]

3. $[R_L, R_N, R_H] \Longrightarrow$ [Training L, No Training, Training H]

4. $[R_L, R_H, R_N] \Longrightarrow$ [Training L, Training H, No Training]

5. $[R_H, R_N, R_L] \Longrightarrow$ [Training H, No Training, Training L]

6. $[R_H, R_L, R_N] \Longrightarrow$ [Training H, Training L, No Training]

Budget constraint not binding If the county has enough funds, such that the budget constraint is not binding ($\lambda = 0$), then the welfare agency provides training to all the individuals for which the (net) returns to training are positive. The allocation between training L and training H is determined by the marginal returns to each pair of treatments compared with the marginal costs of treatment. Hence, inside each skills region the individuals receive the treatment with the greater (net) benefit compared to the other two. The agency chooses the optimal critical values θ_l^u and θ_h^u (where the superindexes stands for "unconstrained") such that

$$
\begin{aligned}
V^m(\theta_l^u) - V^l(\theta_l^u) &= c_m - c_l && (4) \\
V^h(\theta_h^u) - V^m(\theta_h^u) &= c_h - c_m. && (5)
\end{aligned}
$$

The left hand sides of (4) and (5) represent the marginal benefit of one training with respect to the other, while the right hand sides represent the marginal costs. Figure 2 depicts this situation, where panels 2.A to 2.F show each of the possible six cases.

Note that if the individuals do take into account the direct costs of training, c_L and c_H, the solution is the same as would be obtained in the decentralized problem where the individuals are allowed to choose the treatment strategy themselves. For example, individuals could be offered "vouchers", valid to use in any training of their choice, allowing them to keep any difference between the value of the voucher and the cost of the training. In this way, they would completely internalize the cost of training. As such, the model shares the essential features of the Roy Model of self-selection in the labor market (Roy, 1951), in which heterogeneous agents self-assign themselves to occupations according to a principle of comparative advantage.[32]

Budget constraint binding When the budget constraint is binding, using (1) and (2), it is easy to show that the county will choose the θ_l and θ_h that satisfy the following condition

[32] See on the Roy Model Willis (1986) and Heckman and Honore (1990).

(where the superindexes stands for "constrained") :

$$\frac{V^m(\theta_l^c) - V^l(\theta_l^c)}{c_m - c_l} = \frac{V^h(\theta_h^c) - V^m(\theta_h^c)}{c_h - c_m}. \tag{6}$$

That is, the treatment regions are chosen in a way that equalizes the ratio of marginal benefits to the ratio of marginal costs of the treatments across regions. Hence, the county resorts to substitution between treatments until the above condition is satisfied. Although (4) and (5) might appear as similar conditions to (6), it is important to note that the mechanisms that operate in this case are different with respect to the unconstrained case. Specifically, the agency no longer equates just private marginal benefits with private marginal costs, but now has to consider the "social" cost implicit in the fact that providing training to some individuals imply that training has to be denied to other individuals. Equation (6) takes this into account by equating ratios of marginal benefits to marginal costs across the two critical values of θ.

To describe exactly how the agency attains the optimality condition let's consider Case 1 as an example. Faced with a binding budget constraint, the county obviously can train fewer people (with respect to the unconstrained case), the issue is how to decide which individuals to train, and which training to offer them. There are two effects. First, because training H is the most expensive treatment, the county chooses to offer fewer people this treatment. As it is shown in Figure 2.A, the individuals "denied" treatment H (that is, individuals who in the unconstrained case would have received training H), will be offered treatment L. But, as more and more individuals around θ_h^u are offered treatment L instead of treatment H, the lost marginal benefits of training increase (as measured by the vertical distance between the curves $V^H(\theta) - c_H$ and $V^L(\theta) - c_L$). Second, if the county denies training to individuals that otherwise would have received treatment L, starting with the individuals around θ_l^u, the lost marginal benefits are also small at the beginning and increase as more and more people are denied training. The equilibrium will be attained when the vertical distance between $V^H(\theta) - c_H$ and $V^L(\theta) - c_L$, divided by $c_H - c_L$, and between $V^L(\theta) - c_L$ and $V^N(\theta)$, divided by c_L, are equated (which is what equation (6) shows). In the figure the new critical values are represented by θ_l^c and θ_h^c, and the shaded areas roughly indicate the lost "social" benefit of substituting for one treatment or the other.[33]

It is clear from Figure 2.A that total training diminishes (as shown by the translation to the right of θ_l^u to θ_l^c) as well as treatment H (as shown by the translation to the right of θ_l^u to θ_l^c). However, it is not obvious which is the effect on the percentage of people receiving training L, because on one hand it decreases in the lower tail of the distribution of θ, but on

[33]The shaded areas are just an approximation because they do not take into account the division by the cost differences.

the other hand it increases in the upper part of such distribution.[34]

Note that the problem of how the agency assigns participants to training, subject to a common budget constraint, has close parallels to the intrahousehold allocation decisions of parents deciding how much to invest in the education of their heterogeneous-in-skills children.[35] Analogously to the terminology in this literature, the county follows a compensating strategy (i.e. devote more resources to more disadvantaged individuals) when the nature of the training production functions is such that lower skill individuals benefit more from the more expensive training (cases 2, 5 and 6), where as it follows a reinforcing strategy (i.e. devote more resources to more skilled individuals) if the more expensive training favors more skilled individuals (cases 1, 3 and 4). As in the parental allocation literature, even if the county has equal concern for every individual (weights them equally), it will choose not to offer every individual their optimal treatment option, because it takes into account the trade-off generated by the resource constraint.

3.4 Testable implications of earnings-maximization model

The objective of the model is to be able to study how changes in its parameters affect the proportions of people trained (total and in each type of training). These proportions will be denoted in the rest of paper by P_N, P_L, and P_H for treatments N, L and H respectively. Note that $P_N \equiv 1 - P_T$, where $P_T \equiv P_L + P_H$ represents the total proportion of individuals receiving any type of training.

To study the changes in P_N, P_L, and P_H it is clear that the key is to analyze the reaction of the optimal critical values θ_l^c and θ_h^c (where superindexes will be dropped in the remainder of the paper to simplify notation) to changes in the environment faced by the welfare agency.[36] Of interest are the effects of changes in the budget B, the effects of changes in a parameter vector Γ that affects the opportunity cost of training $Y^N(\theta;\Gamma)$, and the effect of changes in a parameter vector Ψ that affects the distribution of skills $F(\theta;\Psi)$.

[34] The intuition behind the other cases is the same, although under Case 2 and Case 4 the proportion of people receiving training 1 unambiguously increases respect to the unconstrained case, while in Case 3 and Case 5 it unambiguously decreases. In Case 6 the change is ambiguous. A formal proof of these results is presented in the next subsection.

[35] See Becker and Tomes (1976), Behrman, Pollak and Taubman (1982) and Behrman (1997). The models of the family in these papers differ from the training problem in that parents seeking to maximize the wealth of their (dynastic) family seek to make decisions about both educational investments and financial transfers (either bequests or inter-vivo transfers), whereas in the assignment to training model agencies only consider how to allocate individuals across types of training.

[36] The two critical values are enough to determine the proportions of each treatment: using the general notation for cases 1 through 6,the proportions are $P_l = F(\theta_l)$, $P_m = F(\theta_h) - F(\theta_l)$, and $P_h = 1 - F(\theta_h)$.

Changes in budget The following proposition indicates the effects of changes in available budget to the county, B, on P_N, P_L and P_H.

Proposition 1 *If the budget B increases then:*
a) the total proportion of people not trained, P_N*, decreases;*
b) the proportion of people receiving training H, P_H*, increases;*
c) the change in the proportion of people receiving training L, P_L*, is ambiguous in Case 1 and Case 6, negative in Case 3 and Case 5, and positive in Case 2 and Case 4.*

Proof. *See Appendix.* ■

The intuition behind this result is the same as the previous subsection. Faced with an increased budget, the county obviously can train more people. The issue is how the agency decides which additional training to offer and to whom. Because training H was offered in all cases to a lower proportion of people than when the budget constraint was not binding, the county increases P_H. This is attained in different ways depending on the case. In Cases 1 and 6 it is done by offering training H to individuals that with the original budget would have been offered training L, in Cases 3 and 5 it is done by offering training H to individuals that otherwise would have been offered treatment N (no training), and in Cases 2 and 4 it is done by a combination of the two previous strategies.

Likewise, the increase in budget makes it profitable for the county to offer more training L than before, in all cases in which the county had substituted training L with treatment N. This effect, plus the increase of P_H in Cases 2 through 5 implies that, unambiguously, P_N always decreases (P_T increases). However, the effects of a budget increase on P_L are not clear: in Cases 2 and 4 P_L unambiguously decreases (because training L was used as substitute for training H), while in Cases 3 and 5 P_L increases (because both training L and H were rationed). In Cases 1 and 6, there are two opposite effects on P_L (as it is shown in the Appendix the total effect will be ambiguous): the positive substitution from no training to training L and the negative substitution from training L to training H.

Changes in the opportunity cost of training To study the effects of changes in the opportunity cost of training, I parameterize Y^N with a vector Γ that measures changes in the earnings generating function under no treatment in two ways: by increasing Y^N in a constant manner, and by increasing it more for more able individuals. Then, $Y^N(\theta;\Gamma) = \gamma_0 + \gamma_1 Y^N(\theta)$, where the vector Γ is formed by two elements: γ_0 which implies a parallel translation of the original function $Y^N(\theta)$ and γ_1 which implies a proportional increase of the returns to skills. Note that a change in Y^N affects also Y^L and Y^H (the earnings under training L and H), but just through the effect of Y^N. This rules out the possibility that changes in the economic situation of the county also affect the treatment effects Δ^L or Δ^H.

The effects of a change in γ_0, a parallel translation in Y^N, are specified in the following proposition, and depend on a condition that is fully explained below.

Proposition 2 *a) In general the effects of a parallel translation in the earnings generating function Y^N are ambiguous;*
b) If the ratio of the opportunity cost to its direct cost of training H is "small enough" relative to the same ratio for training L, then an increase in γ_0 will reduce P_L and will increase P_N and P_H.

Proof. *See Appendix.* ■

This proposition shows that the effect of a change in the opportunity cost of training will have two opposite effects. On one hand, there is a direct effect of increasing the cost of no treatment, which implies that P_N should always increase. How P_L and P_H change is more complex to analyze. Given that the budget is fixed and that P_N increases, with everything else equal, the county has now more budget available for the people that are going to be trained. Following the same logic of Proposition 1, given that training H is always "rationed", there should be a substitution of training L for training H (therefore making P_L decrease and P_H increase). However, because training H is the training that takes longer to be completed, there is an effect by which the opportunity cost of this training with respect to training N increases more than for training L, making training H less desirable. If this effect is big enough, then the substitution will be from training H to training L, which in turn might even imply (if the savings are big enough, given that training H is more expensive) that the increase in P_L makes P_N actually decrease.

Part b) of the proposition states the condition under which this situation will not occur: if the ratio of the opportunity cost to its direct cost of training H is "small enough" with respect to the same ratio for training L unambiguously P_L will decrease, and P_N and P_H will increase. In the Appendix I show that the necessary and sufficient condition for part b) of the proposition is

$$\frac{\tau_H}{c_H} < \frac{\tau_L}{c_L},$$

where the left hand side of the inequality is an expression that captures the ratio of opportunity versus direct cost of training H (because the higher is τ_H, the time that takes to complete the training, the higher is the opportunity cost), and the right hand side captures the same ratio for training L.

As a testable implication, the situation where the "direct effect" of an improvement in the benefits of no treatment dominates the "indirect effect" of changes in the relative benefits of the two types of training, seems more plausible. Therefore, in the empirical analysis, part b) of Proposition 1 will be the one of interest.

The effects of a change in γ_1, a proportional translation in Y^N, are the same as the effects of a change in γ_0. Therefore, in the interest of space, I defer to the Appendix the discussion of the effects of a change in γ_1, presented in Proposition 3, which is analogous to Proposition 2. In summary, the expected effect of an increase in the opportunity cost of training, it that P_L decreases, while P_N and P_H increase.

Changes in the distribution of skills If the distribution of skills of a cohort of individuals to be treated changes, it would be useful to understand how the county would change its optimal decisions. This happens to be a particularly difficult issue to explore. In the Appendix I show, for two distributional assumptions (Uniform and Normal), that the effects of changes in the distribution of skills are ambiguous. The intuition is that depending on the (relative) cost of the treatment towards which the distribution moves, the county might be able to increase the proportion of people receiving that particular treatment, for the cheaper treatments, but not for the expensive treatments. Therefore, changes that would be easy to analyze if the budget constraint were not binding, become extremely complicated with a binding budget constraint. See the Appendix for details.

Summary of testable implications In summary, the testable implications of the models are: an increase in the budget decreases P_N, increases P_H, and has ambiguous effects on P_L (negative in cases 2 and 4, positive in cases 3 and 5 and undeterminate in cases 1 and 6). An increase in the opportunity costs of training increases P_N and P_H and decreases P_L.

3.5 Alternative objective functions

The simple earnings-maximization model assumes that only factors associated with returns to training enter the decision making process of the county welfare agency. However, there are reasons why this might not be the case. Because welfare agencies are public bureaucracies there are many ways in which their decisions will reflect preferences for particular groups or types of training that arise as the interaction of the objectives and preferences of several parties involved.

First, the general oversight of the welfare agency is done by county supervisors, who are elected officers and determine general policy guidelines, as well as appoint the agency managers. As such, the managers preferences should reflect the elected officers preferences. Second, the agency managers and caseworkers might have their own preferences that will affect the agencies decisions.[37] Third, special interest groups (i.e. advocacy groups) might also influence the decisions of the county welfare agency. Therefore, the combination of these factors might generate preferences regarding either individuals or a type of training, that would be reflected on the assignment to training policies.

Political economy models, and the political science literature, suggests that the actions of politicians are influenced by different factors. First, it is well documented in various types of polling data, that citizens often have strong views about people on welfare, the appropriateness of having a welfare system and what types of benefits and services people on welfare should receive (Kluegel and Smith, 1986; Cox and Barrett, 1992; Bean and Papadakis, 1998; Weaver, 2002).[38] Although it is a topic of debate whether politicians and public officers respond to public opinion, or conversely, shape public opinion by pursuing their beliefs, the fact is that there is an interaction between public opinion and policies that can be very important.[39] Second, standard median voter theory suggests that politicians will pursue policies favored by the median voter, which implies that political participation and voting patterns should be an important factor in the politicians actions. Third, because political parties and quasi-public organizations (e.g., unions and special interest groups) often maintain explicit positions, or platforms, relative to the structure of welfare programs, the party affiliation of politicians and/or program administrators might be an important factor in explaining policies.[40]

In what follows I explore two alternative objective functions that, due to some or all the reasons above, reflect specific preferences. One assumes that welfare agencies have preferences for a particular type of training. The other one assumes that the preferences of the agency are over individuals (or groups).

Preferences for one type of training can appear in the particular case of LFA and HCD training, because they are two distinctive approaches with differences that can be deemed philosophical. For example, Hamilton, Freedman et. al. (2001) judge that there has been disagreement over which type of training should be used as the best way of fostering the goal of self sufficiency. They characterize the differences between LFA and HCD as: "[LFA training] emphasizes quick employment, reflecting the belief that individuals can best build their employability and improve their skills, eventually achieving self-sufficiency, through actual

[37]Evidence on the effects of preferences of caseworkers is presented by Heckman, Smith and Taber (1996) regarding acceptance into the JTPA training program.

[38]One theory for the determinants of public support for public programs is self interest, by which individuals that are most likely to benefit from a program favor it (see Cox and Barrett, 1992).

[39]See Jacobs and Shapiro (2002) and Erikson, Mackuen and Stimson (2002) for a debate and some evidence on this issue.

[40]Party identification has been shown to be a good predictor of the attitudes of individuals towards the distribution of services to targeted groups, such as those on welfare (see Kluegel and Smith, 1986 and Cox and Barrett, 1992).

work, even if their initial jobs are minimum wage and without fringe benefits", while "[HCD training] emphasizes initial investments in short-term education and, in some cases, training, reflecting the view that these investments will eventually enable individuals to obtain higher-wage, longer-lasting jobs with health insurance coverage" (Hamilton, Freedman et. al., 2001, p.4). Ideological attitudes towards the role of work and education then can explain preferences for one type of training or the other.[41] These preferences can also reflect heterogeneity in the discount rates of the benefits of training by policy makers, because it has been shown that LFA training benefits are more immediate, while HCD training benefits appear more in the medium and long run (Hotz, Imbens and Klerman, 2000).

To introduce the preferences for one type of training, I consider the case of an agency administrator that does care about the expected returns to the assignment to training policies, but also prefers one type of training to the other. To make notation simpler, suppose that the earnings generating functions are such that, in terms of the prior subsections' terminology, Case 1 is the relevant one.[42] Denoting by α the weight (preference) given to training 1 by the agency administrator, the problem to solve is the following:

$$\max_{\{\theta_l,\theta_h\}} W = \alpha \int_{\theta_l}^{\theta_h} [V^L(\theta) - V^N(\theta) - c_m] dF(\theta) + (1-\alpha) \int_{\theta_h}^{\overline{\theta}} [V^H(\theta) - V^N(\theta) - c_h] dF(\theta) \tag{$P3$}$$

$$s.t.\ [F(\theta_h) - F(\theta_l)]\ c_L + [1 - F(\theta_h)] c_H \leq B.$$

In the Appendix I show that under "moderate" values of α (i.e. such that both types of training are still offered), the empirical implications regarding changes in budget and opportunity costs of training of the simple earnings-maximization model still hold. However, the optimal critical values θ_l and θ_h are different (i.e. the optimal proportions change) and affected by α. The effects of α are clear: higher preference for training L implies a higher proportion of individuals in training L and a lower proportion of no training and training H (see proposition 4 in the Appendix). If the preferences for training L are extreme (α is "large"), then the testable implications of the earnings-maximization model regarding changes in budget and local economics conditions would not hold.

Preferences might be over individuals in a particular position in the skills distributions, and not for training alternatives. In the appendix, I show that this would arise in a specification

[41] The evidence shows that in the specific case of training policies (and attitude towards welfare reform in general) more liberal individuals favor an education-approach, and more conservative individuals favor a work-first approach. See Nathan and Gais (1998) and Weaver (2002).

[42] That is, lower skills individuals do not benefit from training, middle skills individuals benefit from training L, and upper skills individuals benefit from training H.

where the objective function of the program administrator presents inequality aversion.[43] In this case a preference for the training that benefits the lower initial skills group will appear. In the Appendix the first order conditions for a problem in which the program administrator presents inequality aversion are presented, and compared with the ones associated with ($P2$) and ($P3$). It is not difficult to show that for moderate inequality aversion the results would be similar to the ones from solving ($P3$), in which the testable implications regarding budget and opportunity costs of training from the earnings-maximizing model might still hold. However, for high degrees of aversion to inequality the implications would be very different. In general, if the agency administrator (or the caseworker) favors one particular group of individuals[44], and if this group benefits only from a particular training type, then preferences for individuals or for types of training are observationaly equivalent hypotheses.

4 Data

The empirical analysis is based on county level information, and on the aggregation at the county level of individual data on new welfare entrants. The data on training is constructed as quarterly averages from published county level monthly reports by every county of California on the number of people participating in the GAIN and WTW programs, and in each of the activities of the program.[45] Based on these reports, the probabilities of training and participating in activities were constructed.[46] As it was mentioned in section 2, the period under analysis is from quarter 1 of 1994 to quarter 2 of 1999. I use data only for the 25 biggest counties in California, to assure that there is a minimum number of entrants per quarter (at least 100 entrants per quarter).

To measure economic conditions at the county level (as a proxy for the opportunity cost of training), I use published unemployment rates, employment to population ratio (total and in different sectors), and measures of average earnings for all and for specific sectors.[47] In

[43] This type of objective function is proposed by Dehejia (forthcoming).

[44] Heckman, Heinrich and Smith (2002) allow for the possibility in their model that there are preferences for a particular group, not explained by returns to training.

[45] As has been already mentioned in section 2, the data used is from the GAIN25 report (that covers the GAIN Program period and the first year and half of the WTW Program).

[46] The monthly reports record the number of individuals participating in either the GAIN or WTW programs, and the number of "activities" offered during the month. Potentially some individuals might participate in more than one activity, which might imply that more activities are offered than people in the program. Then the proportions of individuals in LFA and HCD programs were calculated by summing up the activities in each category dividing by the sum of LFA and HCD total number of activities, and then multiplying this proportion by the ratio of the number of adults participating in the GAIN/WTW program to the total adult caseload, in the month. In some few cases (for small counties) this last ratio resulted in numbers slightly over one. In these cases the ratios were rounded to one.

[47] The Bureau of Labor and Statistics (BLS) publishes the county level unemployment rates, based on survey methods. The employment and earnings measures are published by the California Employment and Development

addition growth rates of employment and earnings are used.

The budget is approximated by measures of county expenditures on the training programs. These expenditures measure costs associated with providing training services. As it is discussed below, the budget for training is the variable suggested by the theoretical model, but it is very likely to be endogenous to the training policies followed by the counties. Therefore I use an instrumental variables strategy to deal with this problem, as is explained in the next section. The variable to be used as an instrument is the total expenditures associated to running the welfare program (administrative, employment and training services and child care), and do not include the costs of the actual cash aid provided to welfare recipients.[48]

Finally measures of the "political leaning" of the county and of "political participation" are included. These measures are used to take into account that there can be preferences for a type of training or types of individuals by the welfare agencies, which are determined by the party of the agency administrators and/or by the characteristics of the individuals participating in the political process. The political party of the county administrators is not possible to be used because the elections for county supervisors are (at least officially) non-partisan in California (although it is the case that many supervisors do identify with a party). As a proxy then for the "party identification" of the voters in the county, the number of votes cast in the county for the Democratic Party in the State Assembly elections were used to construct the variable "Proportion Votes Democratic Party Assembly". In addition, several measures of the characteristics of the individuals registered to vote in each election are included.[49]

Table 1 presents statewide and yearly averages of the county level data that is used in the regressions.

The individual level data comes from two administrative datasets for the State of California (MEDS and UI base wage files) which provide some demographic and family information and detailed monthly welfare use histories on every individual ever in the welfare system in California, as well as quarterly earnings histories (before, during, and after welfare) for these individuals, as long as their jobs are covered under the Unemployment Insurance system in California (around 90% of the employment of the State).[50]

Department (EDD) and BLS as part of the program known as ES-202 or Covered Employment and Wages.

[48]The data on expenditures comes from the County Financial Analysis Bureau at the California Department of Social Servides (CDSS). For each quarter the data corresponds to the quarter equivalent of the annual expenditures in the corresponding fiscal year.

[49]The data used reflect registration and election results of the general elections of the years 1992, 1994, 1996, and 1998, and comes from the "Statewide Voting Database" of the Institute of Governmental Studies, UC Berkeley. For each quarter, the results from the most recent election were used.

[50]The MEDS (MediCal Eligibility Data System) dataset goes from 1987 to 2000 and contains the welfare use and individual level information (although MEDS identifies welfare cases, not families, it is possible to use the case-level data to construct proxies for measures of family characteristics). The Unemployment Insurance (UI) base wage data goes from 1993 to 2000, and comes from the quarterly reports filed by employers to

Using the MEDS dataset, new entrants to welfare (defined as individuals never on welfare from 1987 on) were identified. As mentioned before, only the 25 biggest counties are analyzed, and also some individuals had to be dropped from the analysis sample because they had missing demographic information, or they belonged to a case which had characteristics that did not allow to construct reliable family structure variables.[51] An additional sample restriction is that only adults 45 years old or younger are analyzed (because with older welfare recipients it is not clear how counties perceive the effectiveness of assigning them to training).

Table 2 presents county and year averages of the characteristics of the sample of entrants to be analyzed. Appendix Table A.2 presents an analysis of the cases dropped from the sample, and the reasons why they were dropped (panel A). It also presents a comparison of the percentages of people dropped from the sample in the 25 counties analyzed versus the rest of the counties (panel B), and it finally shows the distribution of entrants by county (and gives the list of the 25 counties studied) kept in the sample (panel C). In summary, 65% of all the new entrants in the period 1994-1999 are analyzed (449,636 individuals), with a concentration of half of these entrants in some few counties (Los Angeles, Orange, Riverside, San Bernardino, San Diego, and Sacramento).

5 Empirical strategy

The implications derived from the theoretical models in section 3 can be tested empirically, using the county level data on the total proportions of people trained, and receiving LFA and HCD training, and the county level variables presented in the previous section. The following two subsections will detail the econometric specifications to be used.

5.1 Estimation of skills distribution[52]

As the model emphasizes, it is necessary to control for changes in the skills distribution of each cohort of trainees to be able to compare county data across time and counties. Also, having a measure of the skills distribution permits to evaluate empirically the effects of changes in this distribution on the average assignment to training rules of the counties.

---

the California Employment Development Department (EDD). These two datasets were matched through the individuals Social Security Number, to form a dataset that contains not only the (monthly) welfare use history but also the (quarterly) earnings histories of every individual in Welfare in California during the 1990s.

[51] Missing or invalid personal information refers primarily to SSN not valid, and/or missing date of birth, sex or ethnicitcy. Also, because MEDS identifies cases, not families, it is necessary to be careful to identify cases that can be credibly called "families". The cases dropped were all cases in which more than one adult in FG or two adults in UP cases received welfare, had more than two adults not receiving welfare, and had more than 5 kids on welfare.

[52] I want to thank Moshe Buchinsky for suggesting this methodology for estimating the unobserved skills distribution.

The distribution of skills is unobservable for each cohort, at least from the point of view of the econometrician. It is very likely that this is true for the county's welfare agency too, but it is also true that the agency (i.e. the case worker dealing with a particular individual's case) can try to "learn" the value of θ for each individual that enters welfare, by different means. One piece of information is provided by the individual's previous employment and earnings history, as well as personal information (age, civil status, number of kids. age of kids, etc.) that should be very informative regarding the earnings potential of the individual (which is ultimately what θ tries to capture). Thanks to the individual level data available on all the welfare recipients in California, aggregate measures (at the entry cohort-county level) of these variables can be constructed, and used in the analyses.

A second piece of information on the underlying θ for the county is provided by the interaction between the county (the case worker) and the welfare entrant: all counties conduct one-on-one interviews with every entrant to generate an assessment of her potential and weaknesses. In addition, many counties also conduct basic math and language exams to further inform this assessment. As there is no data on the results of these assessments, this is truly unobservable data from the econometric analysis point of view (albeit it is observable for the county). It is also unobservable all other information on the individual that the county observes but is not recorded in the data (most notably educational attainment measures, unfortunately not available in the data sources). Finally there are some other factors (like the individual's motivation) which are not observed by the county (unless the case worker is able to gauge them in the personal interviews) or recorded in any data set.

Let's assume that for an individual i her skill level is $\theta_i = \omega X_i'$, where X_i is a $1 \times K$ vector of individual characteristics observed by the county, and ω is a $1 \times K$ (known) vector of weights used by the county to aggregate this information. Furthermore, assume that X_i is partitioned into two vectors, $X_i = [X_i^o, X_i^u]$, where X_i^o is the vector of characteristics observed in the econometric analysis, and X_i^u is the vector of characteristics that are unobserved in the econometric analysis. Similarly ω can be partitioned as $\omega = [\omega^o, \omega^u]$.

Because the regressions are based on county-level data, it is necessary to introduce measures of the characteristics of the distribution of θ, for example by including in the regression the cohort/county means of the observed variables X^o. However, this has two problems. First, the coefficients associated with these measures will not have a clear interpretation, because they will be implicitly multiplied by ω^o. Second, this strategy leaves out X^u which might include very important factors in the county's decision. Arguably, X^o includes factors like employment and earnings histories that are considered crucial in the selection into training decision, by the training literature[53], but still there are (potentially) other factors which are taken into account

[53] The whole "selection on observables" approach to the non-experimental literature of estimation of training

by the county that are being left out.

The strategy that I follow to deal with this problem is to use the individual level information available on welfare entrants to estimate an index for the left out individual components, and then include features of its distribution in the aggregate level regressions. Specifically, in the tradition of the estimation of earnings equations literature, I calculate earnings equations for the welfare entrants in a period *prior* to entry, and interpret the residuals from these regressions as measures of the skills of the individuals. Subsequently I use measures of the distribution of these residuals in the aggregate level equations.

Because the central interest when calculating the earnings regressions is to obtain the best possible approximation to the skills of the individuals, the strategy for modelling the earnings equation emphasizes prediction of earnings. Running a simple linear equation on the earnings does not seem to be a sensible option, given that (in particular in the earlier years) a large proportion of the individuals were not working before entry into welfare, and then there is a substantial mass at zero earnings. A Tobit model, which would take into account the truncation at zero earnings, might be an option, but the problem is that this model assumes that the same coefficients apply both to the zero part as to the positive part of the earnings distribution.

A better approach is to model both the labor force participation decision, and the earnings generating function, conditional on participation. This is done by using the so called "two-part" model, which first estimates the probability of earnings being positive, and then estimates the earnings equation conditional on positive earnings.[54] That is, the first part estimates a standard probit model

$$\Pr(LFP_i = 1) = \Phi(Z_{1i}\delta_1)$$

where Φ is the CDF of the standard normal distribution, $LFP = 1$ if $Y_i > 0$ (positive earnings) and zero otherwise, and Z_{1i} is a vector of variables that determine the labor force participation decision of individual i. The second part is a linear model

$$\ln(Y_i | LFP_i = 1) = Z_{2i}\delta_2 + u_{2i}$$

where $E(u_{2i}|LFP = 1) = 0$, and Z_{2i} is a vector of variables that determine the (log) earnings of individual i. The vectors Z_{1i} and Z_{2i} include individual demographic variables (language, race, gender), age and age squared, while Z_{1i} also includes family structure variables (number

effects, is based on the assumption that a rich set of variables is available on the individuals, with earnings and employment histories playing a central role.

[54]An alternative would be to use the "Heckit" sample selection model. Leung and Yu (1996) argue that the "two-part" model might be marginally better than the Heckit model for purposes of prediction. Furthermore, both models were estimated with the data, and there are minimum differences between them.

and age groups of kids, entered welfare as single as proxy for civil status), and a dummy that indicates whether the individual received welfare as a child or not. Using this model, each individual's (log) earnings can be predicted as $\ln(\hat{Y}_i) = \Phi(Z_{1i}\hat{\delta}_1) \times Z_{2i}\hat{\delta}_2$.

The measure of skills of an individual i can then be calculated as the residual $\hat{U}_i = \ln(Y_i) - \ln(\hat{Y}_i)$. To make the residuals comparable across counties and cohorts, the above model is estimated pooling all cohorts of entrants together, and including county and cohort fixed effects. This makes it possible to introduce comparable measures of the distribution of the residuals by county and cohort in the county level regressions.

5.2 County level regressions

To test the model presented in section 3, the following three equations are estimated:

$$P_{N,ct} = \beta_{0N} + \beta_{1N}B_{ct} + \beta_{2N}E_{ct} + \beta_{3N}M_{\hat{U},ct} + \beta_{4N}\overline{X}_{ct} + \beta_{5N}Z_{ct} + \eta_c + \tau_t + v_{N,ct} \quad (7)$$

$$P_{L,ct} = \beta_{0L} + \beta_{1L}B_{ct} + \beta_{2L}E_{ct} + \beta_{3L}M_{\hat{U},ct} + \beta_{4L}\overline{X}_{ct} + \beta_{5L}Z_{ct} + \eta_c + \tau_t + v_{L,ct} \quad (8)$$

$$P_{H,ct} = \beta_{0H} + \beta_{1H}B_{ct} + \beta_{2H}E_{ct} + \beta_{3H}M_{\hat{U},ct} + \beta_{4H}\overline{X}_{ct} + \beta_{5H}Z_{ct} + \eta_c + \tau_t + v_{H,ct} \quad (9)$$

where the subindexes c and t refer to county and time, $P_{N,ct}$, $P_{L,ct}$ and $P_{H,ct}$ are respectively the total proportion of people not trained, the proportion of people that received LFA training, and the proportion of people that received HCD training, B_{ct} represents budget, E_{ct} represents measures of local economic conditions, $M_{\hat{U},ct}$ represents moments and percentiles of the distribution of estimated skills distribution of welfare entrants, $\overline{X}_{ct}$ represents average observed characteristics of welfare entrants, and demographic characteristics of the county, and Z_{ct} represents the political variables. The term η_c represents (fixed) unobserved heterogeneity at the county level (such as differences in the administration of the welfare programs at the county level), the term τ_t represents (fixed) time unobserved heterogeneity (such as policy changes at the state level that affect all counties in a period of time), and $v_{N,ct}$, $v_{L,ct}$ and $v_{H,ct}$ are error terms. Note that as $P_N \equiv 1 - P_L + P_H$, all the coefficients in (7) will be equal to one minus the sum of the respective coefficients in (8) and (9). Still, it is interesting to estimate (7) separately because the theoretical model has predictions on the effects on the proportion not trained, even in cases when the testable implications regarding training L are ambiguous.

If $v_{N,ct}$, $v_{L,ct}$ and $v_{H,ct}$ are uncorrelated with all the explanatory variables, this implies that each equation can be consistently estimated, using OLS, allowing for county and time fixed effects (the relevant time period is a fiscal year, the unit of measure of the budget, and it appears as the appropriate variable to capture unobserved changes in policies at the state level).

However, as mentioned in the previous section, the budget measure of resources available for training is derived from the report of expenses of the counties on the training programs. Therefore, is very likely that this variable will be endogenous to the training policies of the counties. To deal with this problem of endogeneity I use an instrumental variables strategy, in which the instrument for the training budget is the total budget (expenditures) of the county. This seems as a reasonable instrument, given that the budget allocations from the State of California to its counties were based, in the period analyzed on a "statewide model" in which the California Department of Social Services allocated funds to counties based on historical spending patterns (for employment and child care services, not for administrative costs).[55]

In the same way, it is necessary to assume that the number and characteristics of the entrants are not affected by the training policies. This seems to be a reasonable assumption, because the entrants to welfare are probably attracted by the cash component of the aid. If they were interested just in training, there are several options available for obtaining training services elsewhere, without the need to apply to welfare. The rest of the county level measures are clearly exogenous to the training policies.

6 Analysis of results

Using the data and the methodology described in the previous two sections, the two steps of the empirical strategy are: 1) estimate the distribution of the skills, and 2) estimate county level regressions based on the training proportions.

6.1 Skills distribution

The two-part model was used to obtain the distribution of skills, based on the earnings of the new entrants to welfare at different periods before welfare. When applying this strategy to data on labor force participation and earnings in quarter 5 through quarter 8 before entry to welfare, the results were very similar. When the same was done using data for quarter 3 and 4 before entry, some noticeable differences appeared, and the results were very different for quarters 1 and 2 before entry. This is consistent with the observed pattern in which, for all cohorts, average earnings are reasonably stable up to quarter 5 before entry, and start to decrease sharply thereafter with a minimum in the quarter of entry to welfare. This is not surprising given that to qualify for welfare the individual (or family in the case of two-parent families) needs to prove that her income is below a certain threshold, and it is a regularity that has been deemed in the literature as the "Ashenfelter dip".

Because the interest is in characterizing the permanent component of the skill measure,

[55]See Hill (2001).

net of transitory shocks that the individuals might be suffering at a particular moment of time (which probably explains the entry to welfare), I do not use data on quarters 1 to 4 before entry. Also, because for the early cohorts fewer quarters of pre-entry data are available, and given that the earnings seem fairly constant for quarters 5 to 8 before entry, I use quarter 5 before entry to obtain the skills distribution.

The estimation was done by pooling together all the cohorts of entrants in all counties, and including time and county fixed effects. Figure 3 shows the (kernel-smoothed) density of the skills distribution (residuals from the two-part model), for selected years (i.e. averaging across counties and quarters within a year).[56]

The figure shows that the distribution of skills experienced a noticeable compression during the period. The biggest movements, though, occur in the upper part of the distribution, with a clear shift toward the center. The effects of these shifts are made clear by Figures 4 to 7. Figure 4 shows that the mean of the distribution has decreased over time, with the differences not statistically significant as is indicated by the confidence intervals. Figure 5, in contrast, shows that the median has increased slightly (around 6%, the skill measure is expressed in logs) until 1998, and remained fairly constant or decreased afterwards, which is a reflection of the movement towards the center of the lower part of the distribution. Figure 6 and 7 show the 10th and 90th percentiles of the skill distribution. While the 10th percentile has remained fairly constant, the 90th percentile experienced a very important change, decreasing more than 20%. This shows that the latest cohorts of entrants were less skilled than the ones at the beginning of the period. This is consistent with the claim by welfare administrators that they had to cope with the individuals hardest to serve in these late periods. This is very consistent with the fact that the economy was very strong through the end of the 90s, meaning that the individuals that really had trouble finding jobs were the ones entering welfare.

6.2 Proportions of training regressions

Table 3 presents the results from running regressions using respectively as dependent variables the proportion of welfare recipients that received LFA training, the proportion of welfare recipients that received HCD training, and the proportion that received no training. Four specifications are shown for each variable, the first and third do not instrument the log(training budget), and the second and fourth use the log(total budget) as an instrument for log(training budget), as it was explained in the previous section. All the specifications include county and fiscal year fixed effects, and the standard errors are robust to heterokedasticity, and corrected

[56] Only years 1994, 1997 and 1999 are presented to make the figure readable, the omitted distributions lie in within the ones displayed.

by county/fiscal year clusters.[57] As it can be seen in columns (1) and (2) of Table 3, the OLS and IV coefficients for the budget measure differ greatly, which highlights the very likely endogeneity of the training budget measure. Note that the table does not show a whole set of variables that were included to control for the demographic characteristics of the entrants to welfare and the adult population in the county.[58]

Effects of changes in the skills distribution The characteristics of the individuals, both observables[59] and unobservables play a role in the determination of the training policies. To characterize the changes in the distribution of skills, the mean, median, 10th and 90th percentile of the skills distribution were included in the regressions.[60] The preferred specification (before including any interaction terms) is the one in Column 2 (that instruments the training budget variable). The decrease of the 90th percentile has a very important effect in increasing the total number of people trained. The same is true for a decrease in the 10th percentile. What is surprising is that the effects of changes in the mean and the median go in opposite directions, but this is a result of the fact that both measures move in opposite directions in the data.

The more interesting question is which of the two types of training is increased when the distribution of skills become more homogenous as result of changes in the upper part of the distribution. As it is explained in the Appendix, the theoretical model gives ambiguous results regarding the expected effects of translations in the distribution, because counties will have to substitute individuals between training programs to satisfy the budget constraint, depending of which training program benefits more the individuals towards which the distribution is moving. The results show that the biggest negative effect of a decrease on the mean of the distribution and increase of the median is on the proportion of people receiving HCD training, which is consistent with a substitution from HCD training to LFA training. The fact that total training decreases with a decrease in the mean reflects the effects of an extra mass of people to train that can not be covered with just substitution between types of training. Also, the coefficient associated to the 90th percentile for LFA training is almost four times the size of the coefficient for HCD training (which also is not statistically significant), showing that, when keeping the amount of resources fixed, the need to provide more training has to be done through the

[57] All the county level regressions were estimated by GMM (which is more efficient in the presence of heteroskedasticity than two stage least squares) and are weighted by average caseload of the counties in the period under analysis. However, for convenience I will still use the term OLS and IV to refer to the regressions.

[58] The variables included in all the regressions are: gender, age and race (black, hispanic, white) composition of the entrants cohorts to welfare, as well as measures of the demographic composition of the adult (18 and plus years old) population in the county (proportions of females, race/ethnicity groups) and log(population).

[59] The coefficients associated to the demographic characteristics of the entrant cohorts and of the county population are not shown for space reasons, they are available from the author upon request.

[60] Higher order moments were also tried, but there were highly correlated with the lower order moments, and they were never significant.

increase of LFA training which is the cheaper training. However, when the decrease in skills comes from the lower part of the distribution (a decrease in the 10th percentile), the effects are more balanced. This, together with the effects of changes in the mean and the median suggest that HCD training is offered to low skills individuals, and LFA training is offered to individuals in the middle part of the skills distribution (Case 6 in the theoretical model).

Note that the above results are holding resources (budget) constant. The results change somehow when relaxing that constraint. In particular Column 4 of Table 3 shows how when the interactions with the resources available to the county is taken into account, the negative effects on HCD training of changes in the distribution are less clear.

Table 4 summarizes the results of Columns 2 and 4 of Table 3 (note that the results in Table 4 are expressed in percentage terms, that is the coefficients in Table 3 are multiplied by 100). Panel A shows the effects on each type of training of a standard deviation increase in all the variables, when no interactions are taken into account. Note that a decrease in the 10th or 90th percentile increase total training, but in particular regarding a decrease in the 90th percentile the increase is much bigger for LFA training (and it is statistically not significant for HCD training). Panel B shows that under a situation in which the available budget is "high" (defined as the mean of the period analyzed plus one standard deviation), the effect of a decrease in the 90th percentile implies a bigger increase in LFA than in HCD training, but that a decrease in the 10th percentile implies only an increase in HCD training, once the availability of resources is taken into account. Under a "low" budget (defined as the mean of the period minus one standard deviation), as expected given the diminished possibility of offering the more expensive training, there is a much smaller effect of the decrease in the 90th percentile with a higher proportional decrease respect to the previous case in HCD training. Also, in the low budget case there is no effect of the changes in the 10th percentile (which are more likely to affect HCD training, if lower skills individuals are receiving this training).

Effects of local economic conditions To capture local economic conditions several variables were tried (not reported in the tables). Unemployment rates had no predictive power, consistent with previous studies that found them as very poor approximations to counties local labor markets (Hoynes, 2000). Also consistent with the same evidence, employment to population ratio and average earnings in the retail trade sector were significant in particular for HCD training and total training. In addition, the quarterly growth rates of employment and earnings in the retail trade sector appear significant for HCD and no training. Other sectors (including overall employment and earnings) were tried, and retail appeared as the one with stronger effects. This should not come as a surprise given that retail is probably one of the more important sectors for low wage workers, and therefore is the sector where welfare recipients would most likely try to obtain jobs.

To evaluate the total effect of changes in the local economic conditions Table 4, panel A shows the effect on the different training programs of an improvement of the economy that implies a higher employment to population ratio, higher earnings, as well as an increase in employment and earnings growth rates. As predicted by the simple earnings-maximization model, the effect on LFA training of an improvement in the economic conditions is negative and the effect on HCD training is positive. The effect on no training also has the expected negative sign.

Effects of budget Under the simple earnings-maximization model, with a budget constraint binding, it is expected that the budget variable will be a fundamental element in determining the training policies. The coefficients in Table 3 (column 2) on the log(training budget) have the predicted signs by the earnings maximization model: they are positive for HCD training and negative for no training. The coefficient on LFA training is ambiguous, as predicted by the model, and although in the regressions they appear as positive, they are statistically not significant.

The size of the coefficients is large, for HCD training the coefficient of 0.19 implies that every 10% of increase in the budget translates roughly in an in increase of two percentage points in the proportion of people receiving HCD training). As a reference point, the average proportion of HCD training in the period was 0.09. In fact Table 4, panel 1, shows how an increase of standard deviation in the budget (respect to the average in the period), implies an increase of 19 percentage points, almost double the actual change in the whole period (from 0.09 to 0.18, see Table 1).

Note in Table 4, panel B, that the effects of the changes in budget do not vary much according to the characteristics of the skills distribution, in particular for HCD training. For LFA training the effect of the budget under a more homogenous distribution (more "compressed") is much bigger, but it remains statistically insignificant.

Effects of political variables Summarizing what has been analyzed up to this point, the results in Tables 3 and 4 are consistent with a simple earnings maximization model: budgets affect positively HCD training and total training, and local economic conditions also affect in the way predicted by the theoretical model. Furthermore, the distribution of skills does appear as an important factor in the training policies, suggesting that HCD training is more responsive to changes in the lower part of the distribution, while LFA training is more responsive to changes in the upper part. This would suggest that a characterization of the results, in terms of the model in Section 3, is that counties are operating under Case 6, that is, offering

HCD training to the lower part of the distribution, LFA training to the middle part of the distribution, and no training to the upper part of the skills distribution.

Although the results in Table 3 seem in general consistent with the simple earnings-maximization model, as it was discussed in Section 3, they are also compatible also with models in which the welfare agency administrators exhibit moderate preferences for a particular type of training or moderate inequality aversion. Therefore, to determine whether the simple earnings-maximization model characterizes the observed training policies, it is necessary to incorporate variables that permit approximate preferences for programs and/or individuals.

Table 5 shows regressions equivalent to the ones in Table 3, but where a whole set of political variables are included.[61] The results strongly suggest that indeed the simple earnings-maximization model is not appropriate, because several of the political variables appear as significant in explaining the training policies.

The "Democratic tendency" of the county, as measured by the proportion of individuals voting for Democratic candidates in the last assembly election, negatively affects LFA training, and positively HCD and total training (although the latter is not statistically significant). Other variables included are registration to vote patterns, and are intended as proxies for political participation of different groups (which might translate as pressure or "mandates" for program administrators to favor particular groups, or types of training). The proportion of adults that are registered as Democrats has no statistically significant effects. However, the variable that has an important effect on HCD training is the proportion of individuals registered as independent or under other political affiliation (not Republican or Democrat). Note that in the period analyzed the proportion of individuals registered in the Republican and Democratic parties decreased, and the Independent/Other category increased (with a stronger negative effect on Democratic registration). It is not surprising that the variable that matters is the registration of Independents/Other because it includes disenchanted Democrats and individuals registered in the Green Party (which has shown an increase in registration in the period).

One of the variables that appears as more consistently affecting training policies is the proportion of adults registered and Hispanic. Its effects are strongly positive on HCD training and total training, and negative on LFA training. This is a period of strong increase in the political participation of Hispanics in California (fueled by several propositions presented to the voters in this period that were perceived as particularly negative for Hispanics),[62] and the

[61] Appendix Table A3 shows the first stage regression, when instrumenting Log (Budget Training) with Log (Total Training), including the effect of political variables (the rest of the variables in this regression are not shown, but all the other variables included in all the regressions were included).

[62] A series of propositions between 1994 and 1998 were passed that were strongly opposed (and generated increased political participation) by Hispanics. Proposition 187 of 1994 (which was later struck down in the

results suggest that when Hispanics have more political clout they strongly favor HCD training. If Hispanics, as a group, perceive that they are likely welfare users, and HCD training is also perceived as a "better quality" training, the results would be consistent with the "self-interest" theory of the determinants of public support for public programs, by which individuals that are most likely to benefit from a program favor it.[63]

Table 5 and 6 presents several specifications allowing for different interactions between political variables, budget and skills distribution. The effects of these specifications are summarized in Table 7 (for the models in Table 5) and in Table 8 (for the models in Table 6).

Table 7, panel A, summarizes the direct effect of the political variables, when not allowing for any interactions. It is interesting that the increase of HCD training associated to voting for Democratic candidates or higher registration of Hispanics and of Independents/Others are through both an increase in total training and a decrease in LFA training. This is when resources (budget) are kept constant. The results change somehow when allowing for different budget environments.

Table 7, panel B, explores the effects of these political variables under two environments: one of high budgets (like the end of the period analyzed), and one of low budgets (similar to the beginning of the period analyzed). The first thing to note is that the effect of the registration of Hispanics and Independent/Others is strong under any budget situation. However, under a high budget the proportion of votes for Democrats has no statistically significant effect on HCD training, but still a negative effect on LFA training. Also, the proportion of registered Democrats appear now as positive on HCD training and on total training (and even a positive but not statistically significant effect on LFA training). The situation is similar under a low budget, but with the proportion of votes for Democratic candidates having a positive and statistically significant effect on HCD training. Under low budget the effects are much smaller which makes sense because it is easier to increase the proportion of HCD training (the more expensive training) when there are more resources available. Still the effects of the political variables are very important.

Table 8 explores the effects of the political variables under different configurations of the skills distribution, regarding the position of the 10th and 90th percentiles. Comparing situations in which the 10th percentile is low or high (defined as the average for the period minus or plus a standard deviation), shows similar results as Table 7 (panel A). The difference between

courts) seeked to deny social services, health care, and public education to undocumented immigrants. Proposition 209 of 1996 banned the use by government agencies of affirmative action programs in employment, and contracting and in university admissions. Proposition 227 of 1998 ended bilingual education in public schools, and it was opposed by some and defended by other Hispanic political leaders (in fact, bilingual education has not been completely eliminated from California public schools through a provision that allowed parents to petition for this type of instruction).

[63]See Cox and Barrett (1992).

these two situations is that the proportion of votes for Democrats, registration of Independents/Other and registration of Hispanics positively affects HCD training more under a low 10th percentile than under a high 10th percentile. That is, when the skill level of the more disadvantaged individuals decrease, more HCD training is offered to these individuals as result of the effect of the political variables. This is consistent with aversion to inequality that gives a higher weight to the more disadvantaged individuals. Regarding the upper part of the skills distribution, two scenarios are studied, with a low and a high 90th percentile (defined also as the average of the period minus or plus a standard deviation). Two results are interesting. First, with a low 90th percentile the proportion of votes for Democratic has a positive effect on LFA training as opposed to the negative or non significant effect in the other cases. Second, the same variable shows an important negative effect on LFA training and a small positive effect on HCD training, in the scenario of a high 90th percentile. Both results seem consistent with a desire for offering some type of training to individuals that are below a certain skill level, and confirm that the individuals that receive LFA training have a higher a skill level than individuals receiving HCD training. Note that in all four scenarios the effect of the registration of Hispanics remains strong and positive.

It is not easy to determine whether the effects of these political variables reflect preferences for a particular type of training (i.e. HCD) or for the individuals that most likely would benefit from that training. The results show that the effects of votes for Democratic candidates are different according to the distribution of skills: when the skills level of the individuals in the upper part of the distribution decrease LFA training increases as much as HCD training with an important increase in total training. In the cases where the upper part of the distribution moves to the right LFA training decreases, and when the lower part of the distribution moves, then HCD training is affected. On the other hand, the effect of the registration of Hispanics appears as important in every case in increasing the proportion of HCD training. That is, Hispanics seem to favor HCD training in general, while the votes for Democratic candidates seem to be associated to more total training in general oriented to individuals below certain skills level. This, unfortunately is not enough to discriminate between the hypotheses of preferences for a type of training or for particular individuals in the skills distribution. What is clear is that the political variables capture preferences that are not consistent with a simple earnings-maximization model.

Effects of political variables including impact on training budget decisions One possibility to explain the effect of the political variables is that they reflect different training budget allocation policies (for example Democrats prefer to spend more on training), but not direct effects on training policies themselves. This can be analyzed in several ways. First, if the political variables affected only training budgets and not the training policies directly, they should not appear as statistically significant in the proportions of training regressions.[64] Second, instead of instrumenting the log(training budget) with log(total budget), the latter can be included directly in the regressions. This changes in the interpretation of the coefficients on the political variables because they will reflect now not only their direct effect on training policies, but also their effect through budget allocation decisions. The fact that except for the registration of Hispanics all other coefficients of the political variables in the first stage regressions (shown in Appendix Table A3) are not statistically significant already suggest that it is not the case that the effect of political variables is only through budget allocation decisions. This is reflected clearly in Table 9 that is equivalent to Table 7 (panel A), showing the effect of changes in political variables on training policies, when log(total budget) is entered directly in the regressions instead of as an instrument for log(training budget). As it is clear of the comparison with Table 7, the only variable that diminishes its effect is the registration of Hispanics. In particular, the results show that the total effect of this variable is of substitution of LFA for HCD training with no statistically significant effect on total training, as a result of the negative effect on training budgets of the registration of Hispanics. What Table 9 makes clear is that the bulk of the effect of the political variables is directly on the training policies, and not through budget allocation decisions.

7 Conclusions

In this paper I modeled the training assignment problem faced by county welfare agencies and considered alternative assumptions about what the agencies seek to maximize when assigning welfare participants to alternative types of training. Using these models I derived empirical implications regarding aggregate training policies, and test them using data for Welfare-to-Work training programs run by California's counties during the 1990s.

My findings imply that a simple earnings-maximization model does not fully characterize the observed assignment decisions of county welfare agencies. The results are consistent with the expected implications of a simple earnings maximization model and show that counties adjust their policies to changes in the environment they face in the way predicted by this model. However, I also find that political variables have a strong effect on the training policies, specifically increasing human capital development training. This is consistent with political

[64] Remember that log(training budget) is being instrumented with log(total budget). It can be proved that if the political variables only affect log(training budget) but not training policies,then if they are included in the first and second stage regressions when they should have been included only in the first stage regressions, their coefficient in the second stage regression should be zero. That is, political variables can be non-zero in the second stage regression only if they affect training poliicies directly, aside from their effect on training budgets.

economy models of assignment to training in which decision makers have moderate preferences for either a particular type of training or a particular group of participants. The results do not allow to differentiate between these two hypotheses, though.

The results from this study are important for several reasons. First, the results raise the possibility that the imposition of work requirements by the federal government after welfare reform might not be consistent with the objectives of the counties. The literature in performance standards shows how standards that do not properly solve the potential tensions between the objectives behind the implementation of the standards and the local decision makers own objectives may cause "cream-skimming", "gaming" of the system, and probably efficiency losses.[65] In this particular case, the fact that the more expensive training (HCD) seems to be directed to low skills individuals, might imply that under the pressure of satisfying performance standards the counties might decrease the provision of training services to the more disadvantaged individuals. This could decrease or increase efficiency depending on whether these individuals are the ones that benefit most from HCD training or not, but clearly would have equity consequences. One could make the argument that the federal government actually is aware of the preferences of the welfare agencies and that the imposition of particular performance standards has the objective of changing the behavior of the counties. That is an open question, but clearly it is necessary for the federal (and state) government to be aware of the nature of the objective functions and preferences of the local decision makers, when designing incentive policies, to understand how agencies might respond to them, and their consequences on the individuals to be trained.

Second, the results are a contribution to the growing literature in treatment assignment, because they show that the often assumed objective of earnings-maximization may be inappropriate. In particular, as future research, it would be interesting to explore whether these results are idiosyncratic to welfare agencies in California, or if they reflect a larger pattern concerning government agencies.

Third, the results of this study are an input for complementary studies of the effects of Welfare to Work programs. The theoretical models suggest that training policies are determined by the characteristics of the pool of individuals to be trained, by exogenous factors to the programs and by the preferences of the program administrators. The empirical evidence supports these implications, which can then be used to learn more about the characteristics of the production functions for training (earnings generating processes). Specifically, Mitnik (2003) uses the variation in the exogenous factors and preference components in training poli-

[65]Evidence on these undesired effects on training programs can be found in Courty and Marschke (1997), Heckman, Heinrich, and Smith (1997 and 2002), and Bell and Orr (2002). Berger, Black and Smith (2000) discuss some similar potential problems associated to statistical profiling methods.

cies to estimate the treatment effects of the Welfare to Work programs, exploiting the fact that different entry cohorts to welfare were subject to different policy regimes.

Finally, a more direct evaluation of the objective function of the welfare agencies would be possible if more detailed information about assignment to training of individuals were available. In particular, using individual level datasets on assignment to training, it would be possible to further explore the nature of the preferences of county administrators, and their implications for the assignment rules. Also, a better understanding of the political processes behind the determination of training policies is necessary to be able to explain the agencies preferences. These will be lines of future research.

Appendix

Second order conditions for (P2) The second order conditions that need to be satisfied around a maximum for P2, denoting by $\mathcal{L}$ the Lagrangian, are:

$$\det(M) \equiv \det \begin{bmatrix} 0 & \frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \lambda} & \frac{\partial^2 \mathcal{L}}{\partial \theta_h \partial \lambda} \\ \frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \lambda} & \frac{\partial^2 \mathcal{L}}{\partial \theta_l^2} & \frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \theta_h} \\ \frac{\partial^2 \mathcal{L}}{\partial \theta_h \partial \lambda} & \frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \theta_h} & \frac{\partial^2 \mathcal{L}}{\partial \theta_h^2} \end{bmatrix} > 0$$

where (valued at the critical values θ_l^* and θ_h^* that solve $(P2)$):
$\frac{\partial^2 \mathcal{L}}{\partial \theta_l^2} = V_\theta^l(\theta_l^*) - V_\theta^m(\theta_l^*)$;
$\frac{\partial^2 \mathcal{L}}{\partial \theta_h^2} = V_\theta^m(\theta_h^*) - V_\theta^h(\theta_h^*)$;
$\frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \theta_h} = 0$;
$\frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \lambda} = f(\theta_l^*)(c_m - c_l)$;
$\frac{\partial^2 \mathcal{L}}{\partial \theta_h \partial \lambda} = f(\theta_h^*)(c_h - c_m)$.

Then, $\det(M) > 0 \Leftrightarrow -(\frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \lambda})^2 \frac{\partial^2 \mathcal{L}}{\partial \theta_h^2} - (\frac{\partial^2 \mathcal{L}}{\partial \theta_h \partial \lambda})^2 \frac{\partial^2 \mathcal{L}}{\partial \theta_l^2} > 0$. Therefore, the sufficient conditions for the critical values that solve $(P2)$ are,

$$f(\theta_l^*)^2(c_m - c_l)^2[V_\theta^h(\theta_h^*) - V_\theta^m(\theta_h^*) > -f(\theta_k^*)^2(c_h - c_m)^2[V_\theta^m(\theta_l^*) - V_\theta^l(\theta_l^*)], \quad \text{(A1)}$$

and will always be satisfied because under Assumption 2 $V_\theta^h(\theta_h) > V_\theta^m(\theta_h)$ and $V_\theta^m(\theta_l) > V_\theta^l(\theta_l)$.■

Necessary conditions for interior solutions to (P2) The following conditions are necessary for an "interior solution" of $(P2)$ in which all three treatments l, m and h (in ascending order of the distribution of θ) are offered:
i) $V^l(\underline{\theta}) - c_l > V^m(\underline{\theta}) - c_m > V^h(\underline{\theta}) - c_h$;
ii) $V^l(\overline{\theta}) - c_l < V^m(\overline{\theta}) - c_m < V^h(\overline{\theta}) - c_h$;
iii) $V^h(\theta_l) - c_h < V^l(\theta_l^*) - c_l$.
iv) $V^h(\theta_l) - c_h < V^m(\theta_l^*) - c_m$.■

Proof of Proposition 1 Equations (6) and (3) form a system of implicit equations

$$\Pi_1(\theta_l, \theta_h, B, \Gamma, \Psi) = \frac{V^h(\theta_h;\Gamma) - V^m(\theta_h;\Gamma)}{c_h - c_m} - \frac{V^m(\theta_l;\Gamma) - V^l(\theta_l;\Gamma)}{c_m - c_l} = 0 \quad \text{(A2)}$$

$$\Pi_2(\theta_l, \theta_h, B, \Gamma, \Psi) = B + F(\theta_l;\Psi)(c_m - c_l) + F(\theta_h;\Psi)(c_h - c_m) - c_h = 0. \quad \text{(A3)}$$

Differentiating (A2) and (A3) with respect to θ_l, θ_h and B, and reordering terms, the following system needs to be solved to obtain the effect of changes of B on θ_l and θ_h:

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} 0 \\ dB \end{bmatrix}, \quad \text{(A4)}$$

where

$$A = \begin{bmatrix} -\frac{V_\theta^m(\theta_l)-V_\theta^l(\theta_l)}{c_m-c_l} & \frac{V_\theta^h(\theta_h)-V_\theta^m(\theta_h)}{c_h-c_m} \\ -f(\theta_l)(c_m-c_l) & -f(\theta_h)(c_h-c_m) \end{bmatrix}. \tag{A5}$$

Note that

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} -f(\theta_h)(c_h-c_m) & -\frac{V_\theta^h(\theta_h)-V_\theta^m(\theta_h)}{c_h-c_m} \\ f(\theta_l)(c_m-c_l) & -\frac{V_\theta^m(\theta_l)-V_\theta^l(\theta_l)}{c_m-c_l} \end{bmatrix} \tag{A6}$$

where

$$\det(A) = \frac{V_\theta^m(\theta_l)-V_\theta^l(\theta_l)}{c_m-c_l} f(\theta_h)(c_h-c_m) + \frac{V_\theta^h(\theta_h)-V_\theta^m(\theta_h)}{c_h-c_m} f(\theta_l)(c_m-c_l). \tag{A7}$$

The sign of this determinant is key to evaluate the effects of changes in B, and it will depend on the particular case under which the county is operating. Because of Assumption 2 it always holds that $V_\theta^m(\theta_l) - V_\theta^l(\theta_l) > 0$ and that $V_\theta^h(\theta_h) - V_\theta^m(\theta_h) > 0$. Therefore, the sign of the determinant will depend on the differences of costs in each case. Given that $c_H > c_L > c_N$, it is easy to see that under Case 1 $c_h > c_m > c_l$; under Case 2 and Case 4 $c_m > c_l$ and $c_h < c_m$; under Case 3 and Case 5 $c_m < c_l$ and $c_h > c_m$; and under Case 6 $c_h < c_m < c_l$. In this way $\det(A)$ will be positive in Cases 1 and 6, and negative in Cases 2, 3, 4, and 5.

Hence, using (A4), (A6) and (A7), the effect of changes in B on the critical values θ_l and θ_h will be:

$$\frac{\partial \theta_l}{\partial B} = -\frac{1}{\det(A)} \frac{V_\theta^h(\theta_h)-V_\theta^m(\theta_h)}{c_h-c_m} \tag{A8}$$

$$\frac{\partial \theta_h}{\partial B} = -\frac{1}{\det(A)} \frac{V_\theta^m(\theta_l)-V_\theta^l(\theta_l)}{c_m-c_l} \tag{A9}$$

where, given the sign of the cost differences and of $\det(A)$ in each case, it is clear that $\frac{\partial \theta_l}{\partial B}$ will be negative in Cases 1, 2 and 4, and it will be positive in Cases 3, 5 and 6. Likewise $\frac{\partial \theta_h}{\partial B}$ will be negative in Cases 1, 3 and 5, and it will be positive in Cases 2, 4 and 6.

Finally, analyzing for each case P_l, P_m, and P_h, the proportion of individuals receiving treatment l, m, and h respectively, it is easy to see that:

$$\frac{\partial P_l}{\partial B} = f(\theta_l)\frac{\partial \theta_l}{\partial B} \tag{A10}$$

$$\frac{\partial P_m}{\partial B} = f(\theta_h)\frac{\partial \theta_h}{\partial B} - f(\theta_l)\frac{\partial \theta_l}{\partial B} \tag{A11}$$

$$\frac{\partial P_h}{\partial B} = -f(\theta_h)\frac{\partial \theta_h}{\partial B}. \tag{A12}$$

Using (A8) and (A9) in (A10), (A11) and (A12), it can be seen then that the proportion of people receiving treatment N will always decrease (i.e. total training will increase) ($\frac{\partial P_N}{\partial B} < 0 \Leftrightarrow \frac{\partial (P_L+P_H)}{\partial B} > 0$), and that the proportion of people receiving training H will always increase ($\frac{\partial P_H}{\partial B} > 0$). The change in the proportion of people receiving training L will be negative in Cases 2 and 4 (where training H is offered to the individuals in the middle of the distribution of θ) and positive in Cases 3 and 5 (where treatment N is offered to the individuals in the middle of the distribution of θ). However, it will ambiguous in Cases 1 and 6 (where treatment L is offered to the individuals in the middle of the distribution of θ), because in the equation $\frac{\partial P_L}{\partial B} = f(\theta_h)\frac{\partial \theta_h}{\partial B} - f(\theta_l)\frac{\partial \theta_l}{\partial B}$ (even without taking into account $f(\theta_l)$ and $f(\theta_h)$), it is ambiguous whether $\frac{\partial \theta_l}{\partial B} > \frac{\partial \theta_h}{\partial B}$ or $\frac{\partial \theta_l}{\partial B} < \frac{\partial \theta_h}{\partial B}$.■

Proof of Proposition 2 Differentiating (A2) and (A3) with respect to θ_l, θ_h and γ_0, and reordering terms, the following system

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} \left(\frac{\kappa_m - \kappa_l}{c_m - c_l} - \frac{\kappa_h - \kappa_m}{c_h - c_m}\right) d\gamma_0 \\ 0 \end{bmatrix}, \tag{A13}$$

where A is defined as in (A5), needs to be solved to obtain the effects of changes of γ_0 on θ_l and θ_h. Note that $\kappa_m - \kappa_l \equiv -(\tau_m - \tau_l)$, $\kappa_h - \kappa_m \equiv -(\tau_h - \tau_m)$, for $l, m, h = N, L, H$, and $\tau_N \equiv 0$.[66] Then, replacing these terms and using (A6), (A7) and (A13), the effects can be expressed as

$$\begin{aligned} \frac{\partial \theta_l}{\partial \gamma_0} &= -\frac{1}{\det(A)} f(\theta_h)(c_h - c_m) \left[\frac{\tau_h - \tau_m}{c_h - c_m} - \frac{\tau_m - \tau_l}{c_m - c_l}\right] & \text{(A14)} \\ \frac{\partial \theta_h}{\partial \gamma_0} &= \frac{1}{\det(A)} f(\theta_l)(c_m - c_l) \left[\frac{\tau_h - \tau_m}{c_h - c_m} - \frac{\tau_m - \tau_l}{c_m - c_l}\right], & \text{(A15)} \end{aligned}$$

where, by the analysis made in the proof of Proposition 1, $\det(A)$ will be positive in Cases 1 and 6 and negative in Cases 2 through 5. It is clear that the sign of the derivatives will depend crucially on the sign of the term between square brackets. Each ratio in the bracketed term represents the ratio of relative opportunity costs of two treatments (the higher τ, the longer the number of periods in which the individual can not generate earnings) versus the relative direct costs of these two treatments (as $c_H > c_L > c_N$ and $\tau_H > \tau_L > \tau_N$, the two ratios will be always positive).

To analyze the changes in the proportions, use expressions analogous to (A10), (A11) and (A12), and introduce (A14) and (A15) to get (rearranging terms):

$$\begin{aligned} \frac{\partial P_l}{\partial \gamma_0} &= -\frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_h - c_m) \left[\frac{\tau_h - \tau_m}{c_h - c_m} - \frac{\tau_m - \tau_l}{c_m - c_l}\right] & \text{(A16)} \\ \frac{\partial P_m}{\partial \gamma_0} &= \frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_h - c_l) \left[\frac{\tau_h - \tau_m}{c_h - c_m} - \frac{\tau_m - \tau_l}{c_m - c_l}\right] & \text{(A17)} \\ \frac{\partial P_h}{\partial \gamma_0} &= -\frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_m - c_l) \left[\frac{\tau_h - \tau_m}{c_h - c_m} - \frac{\tau_m - \tau_l}{c_m - c_l}\right]. & \text{(A18)} \end{aligned}$$

From (A16), (A17) and (A18), it is clear that each proportion change according to a weight given by the difference in direct costs of the other two treatments, times the difference in ratios of opportunity costs. The sign of these differences will be different in each case, but it

[66] In the case in which the discount factor is not one, the terms will be a little bit different, but all the analysis holds the same. Specifically, with $\rho < 1$, $\kappa_m - \kappa_l \equiv -\frac{\rho^{\tau_m} - \rho^{\tau_l}}{1-\rho}$, $\kappa_h - \kappa_m \equiv -\frac{\rho^{\tau_h} - \rho^{\tau_l}}{1-\rho}$.

is straightforward to see that if

$$\frac{\tau_H}{c_H} < \frac{\tau_L}{c_L},$$

the bracketed expression will be negative in Cases 1, 2 and 3, and it will be positive in Cases 4, 5 and 6. This implies that in *every* case $\frac{\partial P_N}{\partial \gamma_0} > 0$, $\frac{\partial P_L}{\partial \gamma_0} < 0$ and $\frac{\partial P_H}{\partial \gamma_0} > 0$.[67] ■

Proportional changes in earnings under no training (proposition 3) The intuition of the results is exactly the same as with changes of γ_0, with the caveat that the condition that assures that the "direct effect" of an increase of the opportunity cost of training is greater than the "indirect effect" has to take into account now the position in the earnings generating function Y^N in which these changes occur (because the change in γ_1 affects more higher skills individuals than lower skills individuals).

Proposition 3 *a) In general the effects of a proportional translation in the earnings generating function Y^N are ambiguous;*
b) If the ratio of the opportunity cost to its direct cost of training H is "small enough" relative to the same ratio for training L, then an increase in γ_1 will reduce P_L and will increase P_N and P_H.

Proof. The sufficient condition for part b) of the proposition is

$$\frac{\tau_H - \tau_L}{c_H - c_L} < \frac{\tau_L}{c_L} \frac{Y^N(\theta_l)}{Y^N(\theta_h)},$$

where the left hand side is the ratio of the differential opportunity cost of training with respect to the differential direct cost, and the right hand side presents the ratio of opportunity to direct cost of training L, adjusted by a number less than 1, the ratio of the earnings function Y^N valued at the optimal values of θ.

To see this, differentiate (A2) and (A3) with respect to θ_l, θ_h and γ_1, and reorder terms, to obtain the system the needs to be solved to obtain the effects of changes of γ_1 on θ_l and θ_h:

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} \left(\frac{\kappa_m - \kappa_l}{c_m - c_l} Y^N(\theta_l) - \frac{\kappa_h - \kappa_m}{c_h - c_m} Y^N(\theta_h) \right) d\gamma_1 \\ 0 \end{bmatrix}, \tag{A19}$$

where A is defined as in (A5). Note that the only difference between (A13) and (A19) is in the term between parentheses in the vector in the right hand side of (A19). However, the interpretation of this term is the same as in Proposition 2, it shows the difference between the ratios of differential opportunity costs of training with respect to the differential direct costs of training. Then analogously to (A16)-(A18), the effects of changes in $\gamma_1$on the proportions

[67] With a discount factor $\rho < 1$, the analogous condition (which has the same economic implications) will be $\frac{1-\rho^{\tau_2}}{c_2} < \frac{1-\rho^{\tau_1}}{c_1}$.

of each treatment are:

$$\begin{aligned}
\frac{\partial P_l}{\partial \gamma_1} &= -\frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_h - c_m) \left[\frac{\tau_h - \tau_m}{c_h - c_m} Y^N(\theta_h) - \frac{\tau_m - \tau_l}{c_m - c_l} Y^N(\theta_l) \right] \\
\frac{\partial P_m}{\partial \gamma_1} &= \frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_h - c_l) \left[\frac{\tau_h - \tau_m}{c_h - c_m} Y^N(\theta_h) - \frac{\tau_m - \tau_l}{c_m - c_l} Y^N(\theta_l) \right] \\
\frac{\partial P_h}{\partial \gamma_1} &= -\frac{1}{\det(A)} f(\theta_l) f(\theta_h)(c_m - c_l) \left[\frac{\tau_h - \tau_m}{c_h - c_m} Y^N(\theta_h) - \frac{\tau_m - \tau_l}{c_m - c_l} Y^N(\theta_l) \right].
\end{aligned}$$

The analysis is the same as in Proposition 2. To assure that the "direct effect" of the increase in γ_1 is greater than the "indirect effect", it is necessary that the following condition be satisfied:

$$\frac{\tau_H - \tau_L}{c_H - c_L} Y^N(\theta_h) < \frac{\tau_L}{c_L} Y^N(\theta_l). \tag{A20}$$

When this condition is satisfied, the bracketed term in the equations above will be negative in Cases 1, 2 and 3, and it will be positive in Cases 4, 5 and 6.[68] This implies that in *every* case $\frac{\partial P_N}{\partial \gamma_1} > 0$, $\frac{\partial P_L}{\partial \gamma_1} < 0$ and $\frac{\partial P_H}{\partial \gamma_1} > 0$. ■

Effects of change in skills distribution Let's call Ψ the vector of parameters that define the distribution of skills, and represent this distribution as $F(\theta; \Psi)$. Using the definitions of the proportions of individuals receiving treatment l, m and h, it is easy to see that the effect on these proportion of a change in Ψ will be composed by two parts: the change in the optimal θ_l and θ_h because of the distributional change, and the change in the mass of the distribution for given θ_l and θ_h:

$$\begin{aligned}
\frac{\partial P_l}{\partial \Psi} &= f(\theta_l) \frac{\partial \theta_l}{\partial \Psi} + F_\Psi(\theta_l) && \text{(A21)} \\
\frac{\partial P_m}{\partial \Psi} &= [f(\theta_h) \frac{\partial \theta_h}{\partial \Psi} + F_\Psi(\theta_h)] - [f(\theta_l) \frac{\partial \theta_l}{\partial \Psi} + F_\Psi(\theta_l)] && \text{(A22)} \\
\frac{\partial P_h}{\partial \Psi} &= -[f(\theta_h) \frac{\partial \theta_h}{\partial \Psi} + F_\Psi(\theta_h)]. && \text{(A23)}
\end{aligned}$$

The effects of a change in Ψ can be analyzed in the same way as with changes in B, γ_0 or γ_1. In particular, differentiating (A2) and (A3) with respect to θ_l, θ_h and Ψ, and reordering terms, the following system

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} 0 \\ [(c_m - c_l) F_\Psi(\theta_l) + (c_h - c_m) F_\Psi(\theta_h)] d\Psi \end{bmatrix}, \tag{A24}$$

where A is defined as in (A5), needs to be solved to obtain the effects of changes of Ψ on θ_l and θ_h. Then, using (A6), (A7), (A8), (A9) and (A20), the effect of changes in Ψ on the critical values θ_l and θ_h can be expressed as:

$$\begin{aligned}
\frac{\partial \theta_l}{\partial \Psi} &= \frac{\partial \theta_l}{\partial B} [(c_m - c_l) F_\Psi(\theta_l) + (c_h - c_m) F_\Psi(\theta_h)] && \text{(A25)} \\
\frac{\partial \theta_h}{\partial \Psi} &= \frac{\partial \theta_h}{\partial B} [(c_m - c_l) F_\Psi(\theta_l) + (c_h - c_m) F_\Psi(\theta_h)]. && \text{(A26)}
\end{aligned}$$

[68] In fact, for Cases 4, 5 and 6, only the condition $\frac{\tau_2}{c_2} < \frac{\tau_1}{c_1}$ is necessary.

The term in the square bracket shows that the interaction of the relative costs of the treatments and the changes in the mass of the distribution of θ, will determine the effect on θ_l and θ_h. However, (A25) and (A26) are not enough to characterize the effects of changes in Ψ on the proportions, because as it is shown in (A22), (A23) and (A24), there will be an extra term, F_Ψ that will make the final effect in most of the cases ambiguous (many of the signs would be ambiguous even without this extra term).

I make two different distributional assumptions. The first is assume that $F(\theta)$ is distributed Uniform $(\underline{\theta}, \overline{\theta})$, and define $\Psi \equiv \overline{\theta}$. This implies studying the effect of the entrance (or exit) of more able individuals (given the properties of the Uniform distribution, an increase in $\overline{\theta}$ imply an increase in both $E(\theta)$ and $V(\theta)$). The results are ambiguous in most cases: $F_{\overline{\theta}}$ is always negative, but $\frac{\partial \theta_l}{\partial \overline{\theta}}$ and $\frac{\partial \theta_h}{\partial \overline{\theta}}$ are some times positive and some times negative (depending on the case). This reflects the fact that depending the (relative) cost of the treatment towards which the distribution moves, the county would be able to increase the proportion of people receiving that particular treatment. The results below show that in the cases (except Case 1, ambiguous) in which training L is offered to individuals of relative less skills than the ones offered training H, P_L decreases unambiguously if $\overline{\theta}$ increases. Intuitively, given that the budget remains constant, one should think that this would imply a less than proportional increase in P_H and an increase in P_N. In the same way, intuition suggests that the opposite effects should occur for cases in which training L is offered to (relatively) more able individuals than individuals offered training H.

The second assumption is that $F(\theta)$ is distributed Normal (μ, σ^2), and $\Psi \equiv [\mu, \sigma^2]$. Below I show that analyzing the effects of changes in μ and in σ^2 gives also ambiguous results. However, because an increase in μ with σ^2 fixed implies that the new distribution stochastically dominates the old one, the intuition would be equivalent to changing $\overline{\theta}$ in the Uniform case.

Uniform case Assume θ is distributed Uniform $(\underline{\theta}, \overline{\theta})$, and define $\Psi = \overline{\theta}$, then $F(\theta) = \frac{\theta - \underline{\theta}}{\overline{\theta} - \underline{\theta}}$, $f(\theta) = \frac{1}{\overline{\theta} - \underline{\theta}}$, which implies that $F_{\overline{\theta}}(\theta) = -\frac{\theta - \underline{\theta}}{(\overline{\theta} - \underline{\theta})^2} = -f(\theta)F(\theta) < 0$ and that $\frac{\partial F_{\overline{\theta}}(\theta)}{\partial \theta} = -f(\theta)^2 < 0$ (that is, an increase in $\overline{\theta}$ implies that the new distribution stochastically dominates the old one, and that the distance between the two cumulative distribution functions increases with θ). The bracketed term in (A21) and (A22) will be negative in Cases 1 and 3, positive in Cases 4 and 6, and will have an ambiguous sign in Cases 2 and 5, which implies that in Cases 1 and 6 $\frac{\partial \theta_l}{\partial \overline{\theta}} > 0$ and $\frac{\partial \theta_h}{\partial \overline{\theta}} > 0$, in Cases 3 and 4 $\frac{\partial \theta_l}{\partial \overline{\theta}} < 0$ and $\frac{\partial \theta_h}{\partial \overline{\theta}} > 0$ and in Cases 2 and 5 the derivatives will have an ambiguous sign. Using (A21), (A22) and (A23), the only unambiguous signs will be $\frac{\partial P_L}{\partial \overline{\theta}} < 0$ in Cases 3 and 4.

Normal case Assume θ is distributed Normal (μ, σ^2), and define $\Psi = [\mu, \sigma^2]$. Then for a particular value of θ, say $\widetilde{\theta}$, $F_\mu(\widetilde{\theta}) = \int_{-\infty}^{\widetilde{\theta}} f(\theta)(\frac{\theta - \mu}{\sigma^2})d\theta = \int_{-\infty}^{\frac{\widetilde{\theta}-\mu}{\sigma}} \phi(\frac{\widetilde{\theta}-\mu}{\sigma})zdz < 0$, where ϕ is the pdf of a standard Normal, and $z \equiv \frac{\theta - \mu}{\sigma}$. Note that although $F_\mu(\widetilde{\theta})$ is negative, it will attain a minimum at $\widetilde{\theta} = \mu$ and then it will be increase asymptotically towards zero with θ. Therefore, when evaluating $F_\mu(\theta_l)$ and $F_\mu(\theta_h)$ in (A21) and (A22), it is not possible to determine which value is higher. However, still an increase in μ implies that the new distribution stochastically dominates the old one, and intuitively results are similar to changes in $\overline{\theta}$ in the Uniform case.

Changes in σ^2 are even more difficult to evaluate because $F_{\sigma^2}(\widetilde{\theta}) = \frac{1}{2\sigma^2}\int_{-\infty}^{\widetilde{\theta}} f(\widetilde{\theta})((\frac{\theta-\mu}{\sigma})^2 - 1)d\theta = \frac{1}{2\sigma^2}[\int_{-\infty}^{\frac{\widetilde{\theta}-\mu}{\sigma}} \phi(\frac{\widetilde{\theta}-\mu}{\sigma})z^2dz - F(\widetilde{\theta})] \lesseqgtr 0$ (it is > 0 when $\widetilde{\theta} < \mu$, $= 0$ when $\widetilde{\theta} = \mu$, and < 0

when $\widetilde{\theta} > \mu$). Hence, it is not possible to evaluate $F_{\sigma^2}(\theta_l)$ versus $F_{\sigma^2}(\theta_h)$ in (A21) and (A22), given that the position of θ_l and θ_h with respect to μ is unknown.■

Preference for one type of training (proposition 4) The solution to $(P3)$ is similar to the solution to $(P2)$. The first order conditions from $(P3)$ are:

$$\theta_l : -\alpha[V^L(\theta_l) - V^N(\theta_l) - c_L]f(\theta_l) + \lambda f(\theta_l)c_L \leq 0 \tag{A27}$$

$$\theta_h : \{\alpha[V^L(\theta_h) - c_L] - (1-\alpha)[V^H(\theta_h) - c_H] + (1-2\alpha)V^N(\theta_h)\}f(\theta_h) + \lambda f(\theta_h)(c_H - c_L) \leq 0 \tag{A28}$$

$$\lambda : B - [F(\theta_h) - F(\theta_l)]c_L - [1 - F(\theta_h)]c_H \geq 0. \tag{A29}$$

Combining (A27) and (A28), and defining $a \equiv \frac{1-\alpha}{\alpha}$ as the preference factor for training H, the agency chooses θ_l and θ_h such that the following condition is satisfied:

$$\frac{V^L(\theta_l) - V^N(\theta_l)}{c_L} = \frac{aV^H(\theta_h) - V^L(\theta_h) - (a-1)[V^N(\theta_h) - c_H]}{c_H - c_L} \tag{A30}$$

That is, as with $(P2)$ the ratio of marginal benefits to marginal costs of each training are equated, but with a weight factor to benefits given by a. It is easy to see that when $\alpha = 1/2$ (i.e. there is equal weight given to each type of training), then $a = 1$, and (A30) collapses to the optimality condition (6).

If α is such that an interior solution occurs in which (A30) is satisfied, then the empirical implications regarding budget and opportunity cost of training under $(P2)$ still hold[69], and empirical implications can be derived for changes in α. For that is necessary that the preference for training L α be in a certain range, as is specified in the following proposition:

Proposition 4 *If the preference for training L is such that* $\alpha < \frac{V^H(\theta_h) - V^N(\theta_h)}{V^H(\theta_h) + V^L(\theta_h) - 2V^N(\theta_h)}$ *then:*
a) the results in Proposition 1 hold;
b) if also $\alpha > \frac{\kappa_N - \kappa_H}{\frac{c_H - c_L}{c_L}\kappa_N - \frac{c_H}{c_L}\kappa_L + \kappa_H}$ *then the results of Proposition 2 hold;*
c) an increase in α *increases* P_L *and decreases* P_N *and* P_H.

Proof. The condition $\alpha < \frac{V^H(\theta_h) - V^N(\theta_h)}{V^H(\theta_h) + V^L(\theta_h) - 2V^N(\theta_h)}$ is a sufficient condition for the second order conditions of $(P3)$ to be satisfied (because it assures that $[aV^H_\theta(\theta_h) - V^L_\theta(\theta_h) - (a-1)V^N_\theta(\theta_h)] > 0$, which is a condition analogous to the one imposed in Assumption 2). Then, (A29) and (A30) can be used to form the following system of equations:

$$\Pi_1(\theta_l, \theta_h, B, \gamma_0) = \frac{V^L(\theta_l) - V^N(\theta_l)}{c_L} - \frac{aV^H(\theta_h) - V^L(\theta_h) - (a-1)[V^N(\theta_h) - c_H]}{c_H - c_L} = 0 \tag{A31}$$

$$\Pi_2(\theta_l, \theta_h, B, \gamma_0) = B - [F(\theta_h) - F(\theta_l)]c_L - [1 - F(\theta_h)]c_H = 0. \tag{A32}$$

Part a) can be proved by differentiating (A31) and (A32) with respect to θ_l, θ_h and B, and

[69] Only the conditions for the results of Proposition 2 to hold, will be specified.

reordering terms to form the following system

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} 0 \\ dB \end{bmatrix}, \tag{A33}$$

where

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} -f(\theta_h)(c_H - c_L) & -[aV_\theta^H(\theta_h) - V_\theta^L(\theta_h) - (a-1)V_\theta^N(\theta_h)] \\ f(\theta_l)c_L & -\frac{V_\theta^L(\theta_l) - V_\theta^N(\theta_l)}{c_L} \end{bmatrix} \tag{A34}$$

and

$$\begin{aligned} \det(A) = {} & \frac{V_\theta^L(\theta_l) - V_\theta^N(\theta_l)}{c_L} f(\theta_h)(c_H - c_L) \\ & + [aV_\theta^H(\theta_h) - V_\theta^L(\theta_h) - (a-1)V_\theta^N(\theta_h)] f(\theta_l) c_L. \end{aligned} \tag{A35}$$

The sign of the determinant in (A35) is positive, and it is straightforward to see that the same results with respect to changes in B as under $(P2)$ hold (see (A8) through (A12)).

Part b) can be proved by differentiating (A31) and (A32) with respect to θ_l, θ_h and γ_0, and using the (A34) and (A35) in the system

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} (\frac{\kappa_L - \kappa_N}{c_L} - \frac{a\kappa_H - \kappa_L - (a-1)\kappa_N}{c_H - c_L}) d\gamma_0 \\ 0 \end{bmatrix} \tag{A36}$$

where under the condition imposed in part b) it can be shown that $\frac{\kappa_L - \kappa_N}{c_L} < \frac{a\kappa_H - \kappa_L - (a-1)\kappa_N}{c_H - c_L}$ and all the results in (A14) through (A18) hold.

Finally, part c) of the proposition can be proved by differentiating (A31) and (A32) with respect to θ_l, θ_h and a, and forming the system

$$\begin{bmatrix} d\theta_l \\ d\theta_h \end{bmatrix} = A^{-1} \begin{bmatrix} \frac{V^H(\theta_h) - V^N(\theta_h) - c_H}{c_H - c_L} da \\ 0 \end{bmatrix} \tag{A37}$$

where (A34) and (A35) still hold. Solving (A37), it is clear that

$$\frac{\partial \theta_l}{\partial a} = \frac{1}{\det(A)} f(\theta_h)(c_H - c_L) \frac{V^H(\theta_h) - V^N(\theta_h) - c_H}{c_H - c_L} \tag{A38}$$

$$\frac{\partial \theta_h}{\partial a} = -\frac{1}{\det(A)} f(\theta_l) c_L \frac{V^H(\theta_h) - V^N(\theta_h) - c_H}{c_H - c_L} \tag{A39}$$

will be positive and negative respectively. This clearly implies that as a increases P_L decreases and P_N and P_H increase (see for example (A10)-(A12)). Given that $a = \frac{1-\alpha}{\alpha}$, then α and a will go in opposite directions, and an increase in α will increase P_L and decrease P_N and P_H. ■

Objective function with inequality aversion The problem in which the program administrator exhibits inequality aversion can be expressed using a Constant Elasticity of Substitution (CES) welfare function for a decision maker that cares about the after treatment earnings distribution and that solves the following problem:

$$\max_{\{\theta_l,\theta_h\}} W = \frac{1}{\varepsilon}\left[\int_{\underline{\theta}}^{\theta_l}[V^l(\theta)]^{\varepsilon}dF(\theta) + \int_{\theta_l}^{\theta_h}[V^m(\theta)]^{\varepsilon}dF(\theta) + \int_{\theta_h}^{\overline{\theta}}[V^h(\theta)]^{\varepsilon}dF(\theta)\right] \tag{$P4$}$$
$$s.t.\ F(\theta_l)c_l + [F(\theta_h) - F(\theta_l)]\ c_m + [1 - F(\theta_h)]c_h \leq B,$$

where ε is the inequality aversion parameter. When $\varepsilon = 1$ there is no inequality aversion, if $\varepsilon = 0$ the utility function becomes logarithmic and implies unitary inequality aversion, and if $\varepsilon \to -\infty$ then the there infinite aversion to inequality and the welfare function becomes rawlsian where only the welfare of the individual worst off matters. Solving for the first order conditions of this problem it is straightforward to show that the optimality condition will be

$$\frac{[V^h(\theta_h)]^{\varepsilon} - [V^m(\theta_h)]^{\varepsilon}}{[V^m(\theta_l)]^{\varepsilon} - [V^l(\theta_l)]^{\varepsilon}} = \frac{c_h - c_m}{c_m - c_l} \tag{A40}$$

which is analogous to the first order condition of $(P2)$, equation (6). In fact $(P2)$ is a special case of $(P4)$ when $\varepsilon = 1$, i.e. there is no aversion to inequality. It can be shown that for some low levels of inequality aversion the results of $(P2)$ will hold, but for higher levels of inequality aversion the results can actually be very different. If the first case occurs, though, it would not be possible to differentiate between preferences towards the lower earnings group introduced by the inequality aversion parameter, from the preferences for a type of training (as in $(P3)$). If the inequality aversion is high, then depending which training is more beneficial for the lower skills individuals, the results from $(P2)$ could be even stronger, or could become completely the opposite. In particular, if training H favors lower skills individuals, the results from $(P4)$ will be similar with respect to changes in budget and local economic conditions to the results from $(P2)$.■

References

[1] Bean, Clive and Elim Papadakis (1998). "A Comparison of Mass Attitudes Towards the Welfare State in Different Institutional Regimes, 1985-1990". International Journal of Public Opinion Research, Vol. 10, Issue 3, (Fall), 211-227.

[2] Becker Gary and Nigel Tomes (1986). "Human capital and the rise and fall of families". Journal of Labor Economics, Vol. 4, S1-S39.

[3] Bell, Stephen and Larry Orr (2002). "Screening (and creaming?) applicants to job training programs: the AFDC homemaker-home health aide demonstrations". Labour Economics 9, 279-301.

[4] Behrman, Jere, Robert Pollak and Paul Taubman (1982). "Parental Preferences and Provision for Progeny". Journal of Political Economy, Volume 90, No. 1, (February), 52-73.

[5] Behrman, Jere (1997). "Intrahousehold Distribution and the Family" in *Handbook of Population and Family Economics*, Volume 1A, Mark Rosenzewig and Oded Stark (ed.), North-Holland.

[6] Berger, Mark, Dan Black, and Jeffrey Smith (2000). "Evaluating Profiling as a Means of Allocating Government Services". Mimeo (May).

[7] Besley, Timothy and Stephen Coate (1992). "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs", American Economic Review, Volume 82, No. 1. (March), 249-261.

[8] Besley, Timothy and Anne Case (2003). "Political Institutions and Policy Choices: Evidence from the United States". Journal of Economic Literature, Volume 41, Issue 1, (March), 7-67.

[9] Cook, Fay and Edith Barrett (1992). *Support for the American Welfare State: the Views of Congress and the Public*, Columbia University Press.

[10] Courty, Pascal and Gerald Marschke (1997). "Measuring Government Performance: Lessons from a Federal Job-Training Program". American Economic Review, Volume 87, Issue 2, 383-388.

[11] Dehejia, Rajeev (forthcoming)."Program Evaluation as a Decision Problem". Journal of Econometrics, (March)..

[12] Dixit, Avinash (2002). "Incentives and Organization in the Public Sector". Journal of Human Resources, Volume 87, Number 4 (Fall), 697-727.

[13] Eberts, Randall, Kevin Hollenbeck, and Joe Stone (2002). "Teacher Performance Incentives and Student Outcomes". Journal of Human Resources, Volume 87, Number 4 (Fall), 778-811.913-927.

[14] Erikson, Roberts (2002). "Panderers or Shirkers: Politicians and Public Policy", in *Navigating Public Opinion: Polls, Policy, and the Future of American Democracy*, J. Manza, F. Cook and B. Page, ed., Oxford University Press.

[15] Hamilton, Gayle, Stephen Freedman, et. al. (2001) "How Effective are Different Welfare-to-Work Approaches? Five Year Adult and Child Impacts for Eleven Programs". National Evaluation of Welfare-to-Work Strategies, MDRC.

[16] Haveman, Robert and Barbara Wolfe (2000). "Welfare to Work in the U.S.: A Model for Other Developed Nations?". International Tax and Public Finance, Volume 7, 95-114.

[17] Heckman, James, Carolyn Heinrich, and Jeffrey Smith (1997). "Assessing the Performance of Performance Standards in Public Bureaucracies". American Economic Review, Volume 87, Issue 2, 389-395.

[18] Heckman, James, Carolyn Heinrich, and Jeffrey Smith (2002). "The Performance of Performance Standards". Journal of Human Resources, Volume 87, Number 4 (Fall), 778-811.

[19] Heckman, James and Bo Honore (1990). "The Empirical Content of the Roy Model". Econometrica, Vol. 58, No. 5, 1121-49

[20] Heckman, James, Robert LaLonde, and Jeffrey Smith (1999). "The Economics and Econometrics of Active Labor Market Programs". in *Handbook of Labor Economics*, Volume 3, Orley Ashenfelter and David Card, ed., Elsevier.

[21] Heckman, James and Richard Robb (1985). "Alternative methods for evaluating the impact of interventions", in *Longitudinal analysis of labor market data,* James Heckman and Burton Singer (ed.), Cambridge University Press.

[22] Heckman, James, Jeffrey Smith and Christopher Taber (1996). "What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program". NBER Working Paper No. 5535, (April).

[23] Hill, Elizabeth (2001). "Improving CalWORKs Program Effectiveness: Changing the Employment Services Budget Process". Legislative Analyst's Office Report, http://www.lao.ca.gov/2001/calworks/021401_calworks.pdf.

[24] Hotz, V. Joseph, Guido Imbens, and Jacob Klerman. "The Long-Term Gains from GAIN: A Re-Analysis of the Impacts of the California GAIN Program". NBER Working Paper No. 8007, (November).

[25] Hoynes, Hillary (2000). "Local Labor Markets and Welfare Spells: Do Demand Conditions Matter?". Review of Economics and Statistics, Volume 82, Number 3, 351-368 (August).

[26] Hoxby, Caroline (2002). "Would School Choice Change the Teaching Profession?". Journal of Human Resources, Volume 87, Number 4 (Fall), 846-891.

[27] Jacobs, Lawrence and Robert Shapiro (2002). "Politics and Policymaking in the Real World: Crafted Talk and the Loss of Democratic Responsiveness", in *Navigating Public Opinion: Polls, Policy, and the Future of American Democracy*, J. Manza, F. Cook and B. Page, ed., Oxford University Press.

[28] Koretz, Daniel (2002). "Limitations in the Use of Achievements Tests as Measures of Educators' Productivity". Journal of Human Resources, Volume 87, Number 4 (Fall), 752-778.

[29] Kluegel, James and Eliot Smith (1986). *Beliefs about Inequality*, Aldine de Gruyer, New York.

[30] Leung, Siu Fai and Shihti Yu (1996). "On the choice between sample selection and two-part models". Journal of Econometrics, Volume 72, Issues 1-2, May-June, 197-229.

[31] McFadden, Daniel (1975)."The Revealed Preferences of a Government Bureaucracy: Theory". The Bell Journal of Economics, Vol. 6. No. 2 (Autumn), 401-416.

[32] McFadden, Daniel (1976)."The Revealed Preferences of a Government Bureaucracy: Empirical Evidence". The Bell Journal of Economics, Vol. 7. No. 1 (Spring), 55-72.

[33] Manski, Charles (2000). "Identification problems and decisions under ambiguity: Empirical analysis of treatment response and normative analysis of treatment choice". Journal of Econometrics 95, 415-442.

[34] Manski, Charles (2001). "Statistical Treatment Rules for Heterogeneous Populations". Mimeo (October).

[35] Manski, Charles, John Newman and John Pepper (2002). "Using Performance Standards to Evaluate Social Programs With Incomplete Outcome Data: General Issues and Application to a Higher Education Block Grant Program". Evaluation Review, Vol. 26, No. 4, (August), 355-355.

[36] Mitnik, Oscar (2003). "Differential Effects of Welfare-to-Work Programs: Identification with Unknown Treatment Status". Mimeo (October).

[37] Moffitt, Robert (1999). "Explaining Welfare Reform: Public Choice and the Labor Market". International Tax and Public Fiance, Volume 6, 289-315.

[38] Moffitt, Robert (2003). "The Temporary Assistance for Needy Families Program" in *Means-Tested Transfer Programs in the United States,* Robert Moffitt (Ed.), National Bureau of Economic Research, The University of Chicago Press.

[39] Nathan, Richard and Thomas Gais (1998). "Implementing the Personal Responsibility Act of 1996: A First Look", Federalism Research Group, The Nelson A. Rockefeller Institute of Government, State University of New York.

[40] Pepper, John (2002)."To Train or Not To Train: Optimal Treatment Assignment Rules Using Welfare-to-Work Experiments". Mimeo (February).

[41] Pepper, John (forthcoming). "Using Experiments to Evaluate Performance Standards: What Do Welfare-to-Work Demonstration Reveal to Welfare Reformers?". Journal of Human Resources.

[42] Person, Torsten, Gerard Rolland and Guido Tabellini (2000). "Comparative Politics and Public Finance". Journal of Political Economy, Volume 108, No. 6, 1121-1161.

[43] Person, Torsten and Guido Tabellini (2002). "Political Institutions and Policy Outcomes: What are the Stylized Facts?". Mimeo, (August).

[44] Roy, A (1951). "Some Thoughts on the Distribution of Earnings". Oxford Economic Papers, Vol. 3, 135-146.

[45] Salamié, Bernard (1997). *The economics of contracts: primer.* MIT Press.

[46] Taylor, John, Ed. (1999). *Monetary Policy Rules*, The University of Chicago Press.

[47] Weaver, R. Kent (2002). "Polls, Priming, and the Politics of Welfare Reform", in *Navigating Public Opinion: Polls, Policy, and the Future of American Democracy*, J. Manza, F. Cook and B. Page, ed., Oxford University Press.

[48] Willis, R. (1986). "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in *Handbook of Labor Economics*, Volume I, O.Ashenfelter and R. Layard, ed., North-Holland.

[49] Wilson, James (1989) *Bureaucracy. What Government Agencies Do and Why They Do It.*Basic Books, Inc.

Table 1. Evolution of county level variables (annual average 25 counties)

Variable	1994	1995	1996	1997	1998	1999 /a
Proportion Not Trained	0.802	0.780	0.753	0.688	0.485	0.287
	(0.10)	(0.13)	(0.15)	(0.19)	(0.26)	(0.22)
Proportion LFA Training	0.108	0.143	0.186	0.247	0.405	0.531
	(0.07)	(0.08)	(0.11)	(0.14)	(0.22)	(0.21)
Proportion HCD Training	0.091	0.077	0.062	0.065	0.110	0.182
	(0.05)	(0.06)	(0.05)	(0.08)	(0.11)	(0.12)
Total Budget ($ millions) /b	35.98	34.96	35.21	33.34	34.93	38.83
	(33.1)	(32.5)	(32.9)	(30.4)	(30.5)	(33.1)
Training Budget ($ millions) /b	5.30	5.18	5.35	6.12	8.79	11.03
	(4.4)	(4.3)	(4.5)	(5.0)	(7.5)	(8.7)
Training Share of Total Budget	0.173	0.175	0.178	0.210	0.272	0.307
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)
Proportion Votes Democratic Party Assembly	0.573	0.498	0.498	0.528	0.528	0.531
	(0.13)	(0.11)	(0.11)	(0.13)	(0.13)	(0.14)
Proportion Adults Registered Democrat	0.315	0.309	0.307	0.307	0.302	0.284
	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Proportion Adults Registered Indep/Other	0.090	0.085	0.089	0.111	0.119	0.112
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Proportion Adults Registered & Hispanic	0.079	0.085	0.084	0.101	0.099	0.102
	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Proportion Adults Registered & Female	0.338	0.333	0.330	0.345	0.339	0.322
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Proportion Adults Registered & 18-34 yrs old	0.212	0.191	0.189	0.193	0.190	0.168
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Proportion Adults Registered & 55+ yrs old	0.128	0.135	0.134	0.144	0.142	0.146
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Proportion Adults Registered to Vote	0.638	0.623	0.623	0.650	0.650	0.607
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)
Proportion Adults Voted in Assembly Election	0.339	0.259	0.257	0.290	0.286	0.294
	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)
Employment/Population Retail	0.104	0.106	0.107	0.108	0.108	0.107
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Average Real Earnings Retail /$1,000	4.844	4.787	4.739	4.846	5.079	5.046
	(0.46)	(0.46)	(0.46)	(0.47)	(0.52)	(0.49)
Quarterly Growth Rate Employment Retail (%)	0.440	0.520	0.300	0.460	0.470	-0.650
	(2.89)	(2.26)	(2.91)	(2.20)	(2.60)	(2.29)
Quarterly Growth Rate Earnings Retail (%)	-0.990	-0.110	-0.210	0.870	1.660	-1.920
	(5.05)	(2.75)	(3.15)	(2.98)	(3.54)	(6.26)
Enrollment Adults in Welfare (AFDC/TANF)	783,679	784,407	751,427	653,562	541,553	465,460

Notes:
Standard deviations between parentheses
/a First two quarters only
/b Budget refers to data on total expenses (not including cash grants) and expenses in training in the welfare program, (AFDC/TANF), where training referst to both GAIN (1994-1997) and WTW (1998-1999) programs.

Table 2. Characteristics new entrants (annual average 25 counties)

Variable	1994	1995	1996	1997	1998	1999 /a
Percentage Female Entrants	0.759	0.758	0.747	0.760	0.751	0.757
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.05)
Average Age at Entry	27.7	27.0	26.6	26.2	25.8	25.6
	(0.91)	(0.84)	(0.84)	(0.82)	(0.95)	(0.91)
Proportion White	0.340	0.350	0.350	0.354	0.335	0.327
	(0.15)	(0.15)	(0.13)	(0.14)	(0.14)	(0.13)
Proportion Hispanic	0.420	0.399	0.394	0.377	0.371	0.371
	(0.16)	(0.16)	(0.15)	(0.15)	(0.15)	(0.14)
Proportion Black	0.150	0.164	0.173	0.194	0.205	0.211
	(0.08)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)
Proportion Other Races	0.090	0.087	0.083	0.075	0.089	0.091
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Average Number of Kids at Entry	1.471	1.390	1.379	1.354	1.330	1.284
	(0.11)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)
Proportion Entrants with Infants at Entry /b	0.216	0.232	0.245	0.266	0.287	0.289
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Proportion Entrants with Toddlers at Entry /c	0.492	0.477	0.491	0.494	0.487	0.473
	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)
Average Quarters Employed 1 Yr before Entry	1.334	1.387	1.415	1.391	1.384	1.458
	(0.17)	(0.15)	(0.16)	(0.14)	(0.13)	(0.13)
Average Quarters Employed 2 Yrs before Entry	-	1.425	1.426	1.387	1.335	1.352
	-	(0.14)	(0.14)	(0.12)	(0.13)	(0.14)
Average Quarters Employed 3 Yrs before Entry	-	0.989	1.280	1.228	1.163	1.160
	-	(0.34)	(0.14)	(0.12)	(0.12)	(0.14)
Average Real Earnings 1 Yr before Entry/$1000	3.842	3.885	3.819	3.560	3.425	3.635
	(0.53)	(0.46)	(0.51)	(0.44)	(0.47)	(0.52)
Average Real Earnings 2 Yrs before Entry/$1000	-	4.819	4.563	4.134	3.740	3.747
	-	(0.55)	(0.54)	(0.47)	(0.53)	(0.62)
Average Real Earnings 3 Yrs before Entry/$1000	-	3.620	4.511	4.018	3.489	3.950
	-	(1.34)	(0.60)	(0.51)	(0.53)	(2.63)

Notes:
Standard deviations between parentheses
/a First two quarters only
/b Infant=less than 1 year old
/c Toddler=1-5 years old

Table 3. Proportions of each type of training as functions of individual variables, local economic conditions and expenditure

Variable	Proportion LFA Training				Proportion HCD Training				Proportion Not Trained			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Log(Training Budget)	0.18***	0.08	0.10	0.32**	-0.06**	0.19***	0.11***	0.28***	-0.08	-0.27**	-0.21*	-0.60***
	(0.05)	(0.10)	(0.11)	(0.15)	(0.03)	(0.06)	(0.04)	(0.05)	(0.05)	(0.13)	(0.11)	(0.15)
Employment/Population Retail Sector	-3.86**	-4.40**	-1.80	-4.77***	3.20***	4.65***	2.59***	3.86***	0.44	-0.25	-1.60	0.92
	(1.73)	(1.75)	(1.62)	(1.74)	(1.17)	(1.19)	(0.86)	(1.02)	(1.94)	(1.87)	(1.66)	(1.75)
Average Earnings Retail Sector	0.00	-0.01	-0.06	0.00	0.18***	0.16***	0.12***	0.17***	-0.18***	-0.16***	-0.13***	-0.17***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)
Quarterly Employment Growth Retail	0.00	0.05	-0.09	0.10	-0.21**	-0.30***	-0.21***	-0.24***	0.15	0.25	0.28	0.14
	(0.17)	(0.18)	(0.15)	(0.16)	(0.09)	(0.09)	(0.08)	(0.09)	(0.20)	(0.21)	(0.18)	(0.19)
Quarterly Avg Earnings Growth Retail	-0.06	-0.04	0.20*	-0.12	-0.35***	-0.34***	-0.38***	-0.40***	0.39***	0.39***	0.27**	0.51***
	(0.11)	(0.11)	(0.11)	(0.12)	(0.06)	(0.06)	(0.06)	(0.06)	(0.11)	(0.12)	(0.12)	(0.13)
Mean Skills Distribution	0.24	0.25*	10.64***	-6.57	0.27***	0.31***	-9.78***	-7.20***	-0.61***	-0.56***	0.55	13.85**
	(0.15)	(0.15)	(4.00)	(5.07)	(0.07)	(0.07)	(1.66)	(2.07)	(0.16)	(0.16)	(4.00)	(5.63)
Median Skills Distribution	0.11	0.11	-12.19***	6.53*	-0.37***	-0.39***	5.10***	6.30***	0.30**	0.27*	1.20	-12.82***
	(0.12)	(0.12)	(2.59)	(3.61)	(0.07)	(0.07)	(1.47)	(1.49)	(0.14)	(0.14)	(2.91)	(4.12)
10th Percentile Skills Distribution	-0.06	-0.07	-2.10*	-0.51	-0.05	-0.05	0.50	1.40*	0.15***	0.12**	1.59	-0.90
	(0.04)	(0.04)	(1.23)	(1.32)	(0.04)	(0.04)	(0.68)	(0.73)	(0.05)	(0.05)	(1.53)	(1.59)
90th Percentile Skills Distribution	-0.23***	-0.23***	-1.90	2.69	-0.04	-0.06	2.88***	2.36***	0.29***	0.29***	-1.44	-5.08**
	(0.06)	(0.06)	(1.66)	(2.08)	(0.04)	(0.04)	(0.61)	(0.72)	(0.07)	(0.07)	(1.73)	(2.14)
Log (Training Budget) * Mean Skills			-0.77***	0.49			0.73***	0.55***			-0.06	-1.04**
			(0.30)	(0.37)			(0.12)	(0.15)			(0.30)	(0.42)
Log (Training Budget) * Median Skills			0.91***	-0.47*			-0.41***	-0.49***			-0.08	0.96***
			(0.19)	(0.27)			(0.11)	(0.11)			(0.22)	(0.31)
Log (Tr. Budget) * 10th Percentile Skills			0.15	0.03			-0.04	-0.11*			-0.11	0.07
			(0.09)	(0.10)			(0.05)	(0.06)			(0.11)	(0.12)
Log (Tr. Budget) * 90th Percentile Skills			0.13	-0.21			-0.21***	-0.18***			0.11	0.39**
			(0.12)	(0.15)			(0.04)	(0.05)			(0.13)	(0.16)
Observations	529	529	529	529	529	529	529	529	529	529	529	529
R-squared	0.80	0.80	0.41	0.81	0.43	0.47	0.51	0.49	0.81	0.83	0.71	0.86
Controls for Demographic Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fiscal Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV for Log(Training Budget)	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
F test budget * skills dist = 0 (p-value)			0.00	0.21			0.00	0.00			0.28	0.03

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
*   significant at 10%; ** significant at 5%; *** significant at 1%

Table 4. Effects of changes in training budget, economic conditions and skills distribution under different scenarios

A. Direct effects (no interactions) (based on models in Table 3, Column 2)

Variable (change)	**LFA Training (x 100)**	**HCD Training (x 100)**	**No Training (x 100)**
Log(Training Budget) (increase 1 std dev)	8.7	19.7 ***	` **
	(10.1)	(5.9)	(13.2)
Local Economic Conditions (increase 1 std dev)	-5.3 ***	11.2 ***	-5.9 ***
	(2.0)	(1.5)	(2.0)
Mean Skills Distribution (decrease 1 std dev)	-1.0 *	-1.3 ***	2.3 ***
	(0.6)	(0.3)	(0.7)
Median Skills Distribution (decrease 1 std dev)	-0.4	1.4 ***	-1.0 *
	(0.4)	(0.3)	(0.5)
10th Percentile Skills Distribution (decrease 1 std dev)	0.4	0.3	-0.7 **
	(0.3)	(0.2)	(0.3)
90th Percentile Skills Distribution (decrease 1 std dev)	3.0 ***	0.7	-3.7 ***
	(0.8)	(0.5)	(0.9)

B. Allowing for interactions between budget and skills distribution (based on models in Table 3, Column 4)

Variable (change)	**LFA Training (x 100)**	**HCD Training (x 100)**	**No Training (x 100)**
Spread skills distribution (low 10th P, high 90th P)			
Log(Training Budget) (increase 1 std dev)	8.1	18.2 ***	-26.2 **
	(9.3)	(5.5)	(11.8)
Compressed skills distribution (high 10th P, low 90th P)			
Log(Training Budget) (increase 1 std dev)	14.2	21.7 ***	-35.9 ***
	(9.2)	(5.0)	(11.4)
High Log(Training Budget)			
Mean Skills Distribution (decrease 1 std dev)	-6.0	-6.9 ***	13.0 ***
	(4.3)	(1.8)	(4.8)
Median Skills Distribution (decrease 1 std dev)	4.2	6.1 ***	-10.2 ***
	(2.8)	(1.1)	(3.2)
10th P Skills Distribution (decrease 1 std dev)	-0.1	1.9 *	-1.8
	(1.6)	(1.0)	(2.0)
90th P Skills Distribution (decrease 1 std dev)	9.3 *	6.4 ***	-15.8 ***
	(5.4)	(2.0)	(5.6)
Low Log(Training Budget)			
Mean Skills Distribution (decrease 1 std dev)	-1.7	-2.2 ***	3.9 ***
	(1.1)	(0.5)	(1.2)
Median Skills Distribution (decrease 1 std dev)	0.5	2.3 ***	-2.8 ***
	(0.8)	(0.4)	(0.9)
10th P Skills Distribution (decrease 1 std dev)	0.3	0.6	-0.9
	(0.5)	(0.4)	(0.6)
90th P Skills Distribution (decrease 1 std dev)	3.6 ***	1.6 **	-5.2 ***
	(1.3)	(0.7)	(1.4)

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%

Results based on coefficients in Table 3 columns 2 and 4. Standard deviation change is with respect to average 1994-1999.

Table 5. Proportions of each type of training adding political variables

Variable	Proportion LFA Training		Proportion HCD Training		Proportion Not Trained	
	(1)	(2)	(1)	(2)	(1)	(2)
Log(Training Budget)	0.09	0.48*	0.17***	-0.11	-0.26**	-0.37
	(0.09)	(0.26)	(0.05)	(0.13)	(0.12)	(0.27)
Proportion Votes Democratic Party	-0.15**	6.55**	0.21***	4.31**	-0.05	-10.80***
	(0.06)	(2.55)	(0.05)	(1.70)	(0.08)	(3.42)
Proportion Adults Registered Democrat	-0.04	-15.11*	-0.33	-11.67***	0.38	27.53***
	(1.06)	(8.54)	(0.62)	(4.45)	(1.14)	(9.15)
Proportion Adults Registered Indep/Other	1.04	-11.16	2.15**	4.91	-3.19*	7.20
	(1.49)	(8.70)	(1.01)	(7.63)	(1.80)	(13.2)
Proportion Adults Registered & Hispanic	-1.94*	21.98**	4.04***	-1.34	-2.09*	-19.88**
	(1.07)	(10.3)	(0.94)	(6.05)	(1.14)	(9.76)
Proportion Adults Registered & Female	-1.37	5.62	1.49	-11.17	-0.08	3.40
	(2.27)	(18.1)	(1.38)	(12.6)	(2.35)	(21.9)
Proportion Adults Registered & 18-34 yrs old	2.64***	20.27*	2.66***	-0.21	-5.30***	-18.88
	(0.88)	(11.9)	(0.87)	(7.46)	(1.10)	(11.8)
Proportion Adults Registered & 55+ yrs old	1.21	11.45	3.13**	31.07***	-4.35**	-40.98*
	(1.46)	(15.7)	(1.25)	(10.3)	(1.85)	(21.4)
Proportion Adults Registered to Vote	-0.42	1.81	-3.62***	-6.73***	4.01**	4.77**
	(1.69)	(1.91)	(1.09)	(1.14)	(1.85)	(2.21)
Proportion Adults Voted in Election	0.08	-7.10**	-0.55***	-0.43	0.46***	7.41**
	(0.16)	(3.20)	(0.12)	(1.84)	(0.15)	(3.46)
Mean Skills Distribution	0.21	0.22	0.25***	0.23***	-0.46***	-0.46***
	(0.15)	(0.15)	(0.08)	(0.08)	(0.17)	(0.17)
Median Skills Distribution	0.13	0.10	-0.34***	-0.22***	0.21	0.12
	(0.12)	(0.14)	(0.07)	(0.07)	(0.14)	(0.17)
10th Percentile Skills Distribution	-0.07	-0.07*	-0.07*	-0.06*	0.13**	0.14***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)
90th Percentile Skills Distribution	-0.20***	-0.20***	-0.06	-0.05	0.26***	0.26***
	(0.06)	(0.06)	(0.04)	(0.04)	(0.07)	(0.07)
Log(Tr. Budget) * Prop. Votes Democratic Party		-0.47***		-0.29**		0.76***
		(0.18)		(0.12)		(0.24)
Log(Tr. Budget) * Prop. Ad. Regist. Democrat		1.09*		0.93***		-2.07***
		(0.62)		(0.33)		(0.67)
Log(Tr. Budget) * Prop. Ad. Regist. Indep/Other		0.73		-0.09		-0.70
		(0.63)		(0.55)		(0.93)
Log(Tr. Budget) * Prop. Ad. Regist. & Hispanic		-1.67**		0.40		1.22*
		(0.76)		(0.43)		(0.72)
Log(Tr. Budget) * Prop. Ad. Regist. & Female		-0.65		1.18		-0.37
		(1.32)		(0.92)		(1.63)
Log(Tr. Budget) * Prop. Regist. & 18-34 yrs old		-1.34		0.20		1.06
		(0.83)		(0.51)		(0.81)
Log(Tr. Budget) * Prop. Regist. & 55+ yrs old		-0.80		-2.14***		2.83*
		(1.20)		(0.78)		(1.65)
Log(Tr. Budget) * Prop. Ad. Voted in Election		0.53**		-0.02		-0.50**
		(0.23)		(0.13)		(0.25)
Observations	529	529	529	529	529	529
R-squared	0.80	0.82	0.49	0.53	0.83	0.87
Controls for Local Economic Conditions	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Demographic Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Fiscal Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
IV for Log(Training Budget)	Yes	Yes	Yes	Yes	Yes	Yes
F test pol var / interactions= 0 (p-value)	0.00	0.01	0.00	0.00	0.00	0.00

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6. Proportions of each type of training including interactions of political variables

Variable	Proportion LFA Training	Proportion HCD Training	Proportion Not Trained
	(1)	(1)	(1)
Log(Training Budget)	0.12	0.15***	-0.29**
	(0.09)	(0.04)	(0.11)
Proportion Votes Democratic Party	3.44***	0.81*	-4.22***
	(0.68)	(0.42)	(0.69)
Proportion Adults Registered Democrat	-9.92***	-3.19**	12.40***
	(2.73)	(1.57)	(3.00)
Proportion Adults Registered Indep/Other	-1.22	4.30*	-4.37
	(3.95)	(2.60)	(3.98)
Proportion Adults Registered & Hispanic	-3.67*	3.61**	0.61
	(2.20)	(1.48)	(2.19)
Proportion Adults Registered & Female	-0.64	2.94	-4.56
	(7.75)	(4.60)	(7.67)
Proportion Adults Registered & 18-34 yrs old	12.91***	-4.99***	-7.86***
	(2.70)	(1.62)	(2.46)
Proportion Adults Registered & 55+ yrs old	0.72	0.15	-1.53
	(3.14)	(1.48)	(3.50)
Proportion Adults Registered to Vote	0.00	-1.10	2.92
	(5.20)	(2.89)	(5.07)
Proportion Adults Voted in Election	-3.80***	0.91	3.15***
	(0.87)	(0.62)	(1.04)
Mean Skills Distribution	4.58	-6.59***	1.93
	(3.05)	(2.07)	(2.79)
Median Skills Distribution	-2.52	0.67	1.93
	(1.97)	(1.03)	(1.81)
10th Percentile Skills Distribution	-1.42	2.15***	-0.76
	(1.02)	(0.77)	(0.95)
90th Percentile Skills Distribution	-1.31	1.50**	-0.12
	(0.83)	(0.63)	(0.78)
Prop. Votes Democratic Party * Mean Skills	7.96***	4.95***	-13.05***
	(1.97)	(1.68)	(2.25)
Prop. Ad. Regist. Democrat * Mean Skills	-18.87***	-12.11***	30.42***
	(6.64)	(4.41)	(7.65)
Prop. Ad. Regist. Indep/Other * Mean Skills	-30.30**	2.13	28.12*
	(14.1)	(8.57)	(14.6)
Prop. Ad. Regist. & Hispanic * Mean Skills	-12.42**	11.96***	1.40
	(5.65)	(4.37)	(5.58)
Prop. Ad. Regist. & Female * Mean Skills	-33.86	23.34	10.95
	(28.3)	(17.7)	(32.1)
Prop. Regist. & 18-34 yrs old * Mean Skills	14.10***	-20.01***	6.18
	(4.93)	(4.12)	(5.76)
Prop. Regist. & 55+ yrs old * Mean Skills	-0.17	0.44	-0.24
	(6.43)	(3.19)	(7.19)
Prop Ad. Registered to Vote * Mean Skills	21.27	2.93	-24.07
	(17.4)	(10.7)	(17.8)
Prop. Ad. Voted in Election * Mean Skills	-8.57***	0.80	7.69***
	(2.50)	(1.52)	(2.70)
Prop. Votes Democratic Party * Median Skills	-4.03**	-4.07***	8.19***
	(1.64)	(0.77)	(1.68)
Prop. Ad. Regist. Democrat * Median Skills	4.69	13.90***	-18.18***
	(4.71)	(2.35)	(5.18)
Prop. Ad. Regist. Indep/Other * Median Skills	54.77***	-4.68	-50.99***
	(9.89)	(6.01)	(10.6)
Prop. Ad. Regist. & Hispanic * Median Skills	7.73	2.90	-11.36**
	(4.75)	(2.96)	(4.55)
Prop. Ad. Regist. & Female * Median Skills	88.65***	-7.64	-83.71***
	(17.0)	(15.1)	(21.2)
Prop. Regist. & 18-34 yrs old * Median Skills	0.41	6.45**	-6.57
	(5.10)	(3.14)	(4.89)
Prop. Regist. & 55+ yrs old * Median Skills	4.23	-5.02*	-0.01
	(5.04)	(2.66)	(5.01)
Prop Ad. Registered to Vote * Median Skills	-55.18***	-0.22	56.71***
	(10.7)	(7.95)	(12.7)
Prop. Ad. Voted in Election * Median Skills	3.78***	-1.71*	-1.84
	(1.37)	(0.98)	(1.67)

(continues next page)

Table 6. Proportions of each type of training including interactions of political variables (continuation)

Variable	Proportion LFA Training	Proportion HCD Training	Proportion Not Trained
	(1)	(1)	(1)
Prop. Votes Democratic Party * 10th P Skills	-1.01	-0.43	1.40*
	(0.70)	(0.55)	(0.78)
Prop. Ad. Regist. Democrat * 10th P Skills	4.40*	-4.47***	0.33
	(2.41)	(1.60)	(2.64)
Prop. Ad. Regist. Indep/Other * 10th P Skills	6.50	-4.03	-2.50
	(4.38)	(3.00)	(5.10)
Prop. Ad. Regist. & Hispanic * 10th P Skills	4.03*	-5.27***	1.36
	(2.16)	(1.73)	(2.03)
Prop. Ad. Regist. & Female * 10th P Skills	-2.55	6.95	-4.86
	(10.1)	(6.37)	(10.6)
Prop. Regist. & 18-34 yrs old * 10th P Skills	1.42	-1.32	-0.45
	(2.37)	(2.11)	(2.35)
Prop. Regist. & 55+ yrs old * 10th P Skills	0.82	-3.00**	1.81
	(2.13)	(1.48)	(2.78)
Prop Ad. Registered to Vote * 10th P Skills	0.21	-2.63	2.69
	(6.58)	(4.02)	(6.31)
Prop. Ad. Voted in Election * 10th P Skills	-0.98	1.52*	-0.29
	(1.30)	(0.78)	(1.64)
Prop. Votes Democratic Party * 90th P Skills	-4.23***	-0.95	5.11***
	(0.70)	(0.60)	(0.81)
Prop. Ad. Regist. Democrat * 90th P Skills	12.19***	0.65	-12.22***
	(2.82)	(1.63)	(3.09)
Prop. Ad. Regist. Indep/Other * 90th P Skills	9.13*	-5.36*	-2.76
	(4.80)	(2.86)	(5.09)
Prop. Ad. Regist. & Hispanic * 90th P Skills	4.78***	-3.94***	-1.13
	(1.71)	(1.45)	(1.62)
Prop. Ad. Regist. & Female * 90th P Skills	2.22	4.13	-5.19
	(10.4)	(6.36)	(11.2)
Prop. Regist. & 18-34 yrs old * 90th P Skills	-6.94***	4.49**	2.24
	(2.01)	(1.80)	(2.07)
Prop. Regist. & 55+ yrs old * 90th P Skills	1.13	-1.02	0.29
	(2.40)	(1.07)	(2.77)
Prop Ad. Registered to Vote * 90th P Skills	-3.91	-3.43	6.22
	(6.65)	(3.83)	(6.38)
Prop. Ad. Voted in Election * 90th P Skills	2.79***	-0.36	-2.46***
	(0.66)	(0.55)	(0.82)
Observations	529	529	529
R-squared	0.84	0.58	0.88
Controls for Local Economic Conditions	Yes	Yes	Yes
Controls for Demographic Characteristics	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes
Fiscal Year Fixed Effects	Yes	Yes	Yes
IV for Log (Training Budget)	Yes	Yes	Yes
F test pol var / interactions= 0 (p-value)	0.00	0.00	0.00

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7. Effects of changes in political variables and interactions with training budget
(based on models in Tables 5)

A. Direct effects (no interactions) (based on models in Table 5, Column 1)

Variable (change)	**LFA Training (x 100)**	**HCD Training (x 100)**	**No Training (x 100)**
Log(Training Budget) (increase 1 std dev)	9.7	18.2 ***	-27.8 **
	(9.6)	(5.6)	(12.7)
Proportion Votes Democratic Party (increase 1 std dev)	-2.0 **	2.6 ***	-0.7
	(0.8)	(0.6)	(1.0)
Proportion Adults Registered Democrat (increase 1 std dev)	-0.2	-1.8	2.0
	(5.7)	(3.3)	(6.2)
Proportion Adults Registered Indep/Other (increase 1 std dev)	2.7	5.7 **	-8.4 *
	(3.9)	(2.6)	(4.7)
Proportion Adults Registered & Hispanic (increase 1 std dev)	-5.4 *	11.2 ***	-5.8 *
	(3.0)	(2.6)	(3.2)
Proportion Adults Registered & 18-34 yrs old (increase 1 std dev)	8.6 ***	8.7 ***	-17.3 ***
	(2.9)	(2.8)	(3.6)
Proportion Adults Registered & 55+ yrs old (increase 1 std dev)	5.1	13.1 **	-18.3 **
	(6.1)	(5.2)	(7.8)

B. Allowing for interactions of political variables with training budget
(based on models in Table 5, Column 2)

Variable (change)	**LFA Training (x 100)**	**HCD Training (x 100)**	**No Training (x 100)**
High Log(Training Budget)			
Prop. Votes Democratic Party (increase 1 std dev)	-14.7 ***	-5.3	19.9 ***
	(5.0)	(3.3)	(6.7)
Prop. Adults Registered Democrat (increase 1 std dev)	14.0	18.9 ***	-33.3 ***
	(10.8)	(6.2)	(11.6)
Prop. Adults Registered Indep/Other (increase 1 std dev)	1.8	9.1 **	-11.2
	(6.4)	(4.3)	(7.7)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-14.3 **	14.1 ***	-0.2
	(6.7)	(4.0)	(6.8)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	-5.0	9.8 **	-5.3
	(6.4)	(4.1)	(6.6)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	-6.6	-15.5	21.4
	(17.7)	(11.0)	(24.6)
Low Log(Training Budget)			
Prop. Votes Democratic Party (increase 1 std dev)	-2.0 **	2.5 ***	-0.5
	(0.8)	(0.6)	(0.9)
Prop. Adults Registered Democrat (increase 1 std dev)	1.6	8.3 **	-9.8
	(6.5)	(3.3)	(6.8)
Prop. Adults Registered Indep/Other (increase 1 std dev)	-2.2	9.6 ***	-7.3
	(4.6)	(2.6)	(5.4)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-4.6	11.8 ***	-7.3 *
	(3.5)	(2.6)	(4.0)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	4.3	8.5 ***	-12.6 ***
	(3.8)	(2.9)	(4.8)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	0.5	3.4	-3.6
	(9.5)	(5.4)	(12.2)

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%

Results based on coefficients in Table 5, Columns 1 and 2. Standard deviation change is with respect to average 1994-1999.

Table 8. Effects of changes in political variables interacted with changes in skills distributior
(based on models in Table 6, Column 1)

Variable (change)	LFA Training (x 100)	HCD Training (x 100)	No Training (x 100)
Low 10th Percentile			
Prop. Votes Democratic Party (increase 1 std dev)	-0.5	3.9 ***	-3.4 ***
	(1.0)	(0.7)	(1.0)
Prop. Adults Registered Democrat (increase 1 std dev)	-3.2	-0.2	1.9
	(6.6)	(3.1)	(6.4)
Prop. Adults Registered Indep/Other (increase 1 std dev)	2.9	4.8 **	-8.2 **
	(4.0)	(2.2)	(4.1)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-5.1	8.2 ***	-2.5
	(3.5)	(2.2)	(3.0)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	15.4 ***	1.2	-16.3 ***
	(3.7)	(2.7)	(4.0)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	4.6	6.0	-10.4
	(7.0)	(4.3)	(7.8)
High 10th Percentile			
Prop. Votes Democratic Party (increase 1 std dev)	-2.0 *	3.2 ***	-1.3
	(1.2)	(0.7)	(1.2)
Prop. Adults Registered Democrat (increase 1 std dev)	-0.4	-3.1	2.2
	(6.9)	(3.3)	(6.6)
Prop. Adults Registered Indep/Other (increase 1 std dev)	4.9	3.6	-9.0 **
	(4.1)	(2.4)	(4.5)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-3.8	6.5 ***	-2.1
	(3.7)	(2.2)	(3.2)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	15.9 ***	0.6	-16.5 ***
	(3.8)	(3.0)	(4.0)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	5.0	4.5	-9.5
	(6.9)	(4.4)	(7.5)
Low 90th Percentile			
Prop. Votes Democratic Party (increase 1 std dev)	5.8 ***	5.2 ***	-10.8 ***
	(1.5)	(1.2)	(1.7)
Prop. Adults Registered Democrat (increase 1 std dev)	-10.4	-2.1	10.6
	(7.0)	(3.5)	(7.0)
Prop. Adults Registered Indep/Other (increase 1 std dev)	0.7	6.0 **	-7.6 *
	(4.3)	(2.4)	(4.0)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-6.2 *	8.8 ***	-1.9
	(3.7)	(2.4)	(3.2)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	18.6 ***	-1.0	-17.3 ***
	(4.0)	(2.7)	(4.1)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	4.2	5.8	-10.1
	(7.3)	(4.4)	(8.2)
High 90th Percentile			
Prop. Votes Democratic Party (increase 1 std dev)	-8.3 ***	2.0 *	6.2 ***
	(1.5)	(1.1)	(1.6)
Prop. Adults Registered Democrat (increase 1 std dev)	6.8	-1.2	-6.5
	(7.0)	(3.2)	(6.6)
Prop. Adults Registered Indep/Other (increase 1 std dev)	7.0	2.4	-9.5 *
	(4.4)	(2.5)	(5.1)
Prop. Adults Registered & Hispanic (increase 1 std dev)	-2.7	5.9 ***	-2.7
	(3.6)	(2.1)	(3.1)
Prop. Adults Regist. & 18-34 yrs old (increase 1 std dev)	12.7 ***	2.8	-15.4 ***
	(3.6)	(3.1)	(4.0)
Prop. Adults Regist. & 55+ yrs old (increase 1 std dev)	5.4	4.7	-9.8
	(6.7)	(4.4)	(7.3)

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
*   significant at 10%; ** significant at 5%; *** significant at 1%

Results based on coefficients in Table 6, Column 1. Standard deviation change is with respect to average 1994-1999.

Table 9. Total effects of changes in political variables (including effects on budget allocations)

Variable (change)	LFA Training (x 100)	HCD Training (x 100)	No Training (x 100)
Log(Total Budget) (increase 1 std dev)	10.3	19.3 ***	-29.4 **
	(9.8)	(5.8)	(12.8)
Proportion Votes Democratic Party (increase 1 std dev)	-2.0 ***	2.6 ***	-0.6
	(0.7)	(0.7)	(1.1)
Proportion Adults Registered Democrat (increase 1 std dev)	-0.6	-2.6	3.2
	(5.5)	(3.5)	(6.2)
Proportion Adults Registered Indep/Other (increase 1 std dev)	2.6	5.5 **	-8.1 *
	(3.7)	(2.7)	(4.6)
Proportion Adults Registered & Hispanic (increase 1 std dev)	-6.8 ***	8.4 ***	-1.6
	(2.6)	(2.5)	(2.9)
Proportion Adults Registered & 18-34 yrs old (increase 1 std dev)	8.6 ***	8.7 ***	-17.3 ***
	(2.7)	(2.8)	(3.3)
Proportion Adults Registered & 55+ yrs old (increase 1 std dev)	4.6	12.5 **	-17.1 **
	(6.1)	(5.2)	(7.7)

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1
Training Policies in California
(Average 25 Counties)

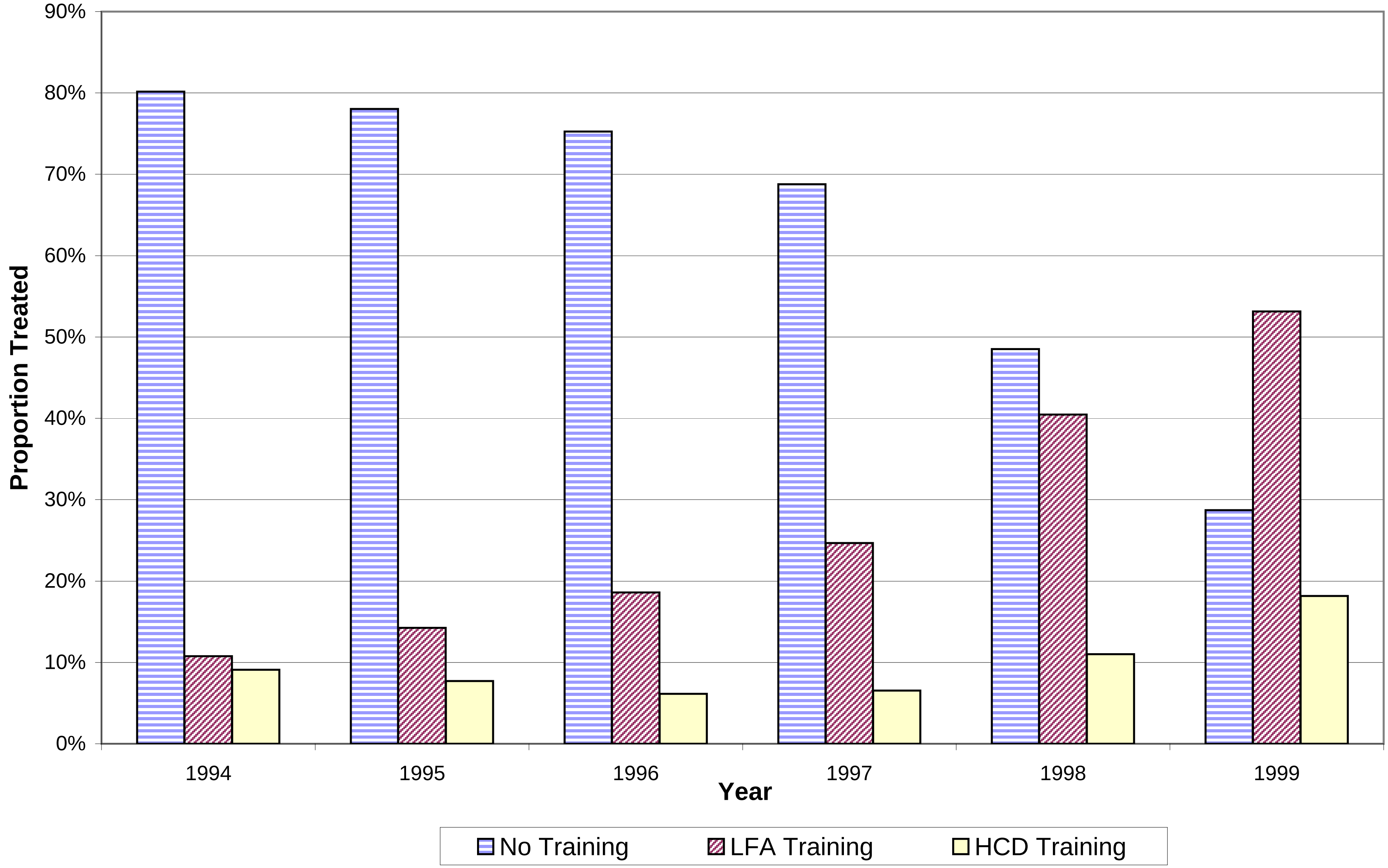


Figure 2. Interior Solution Cases for (P2)

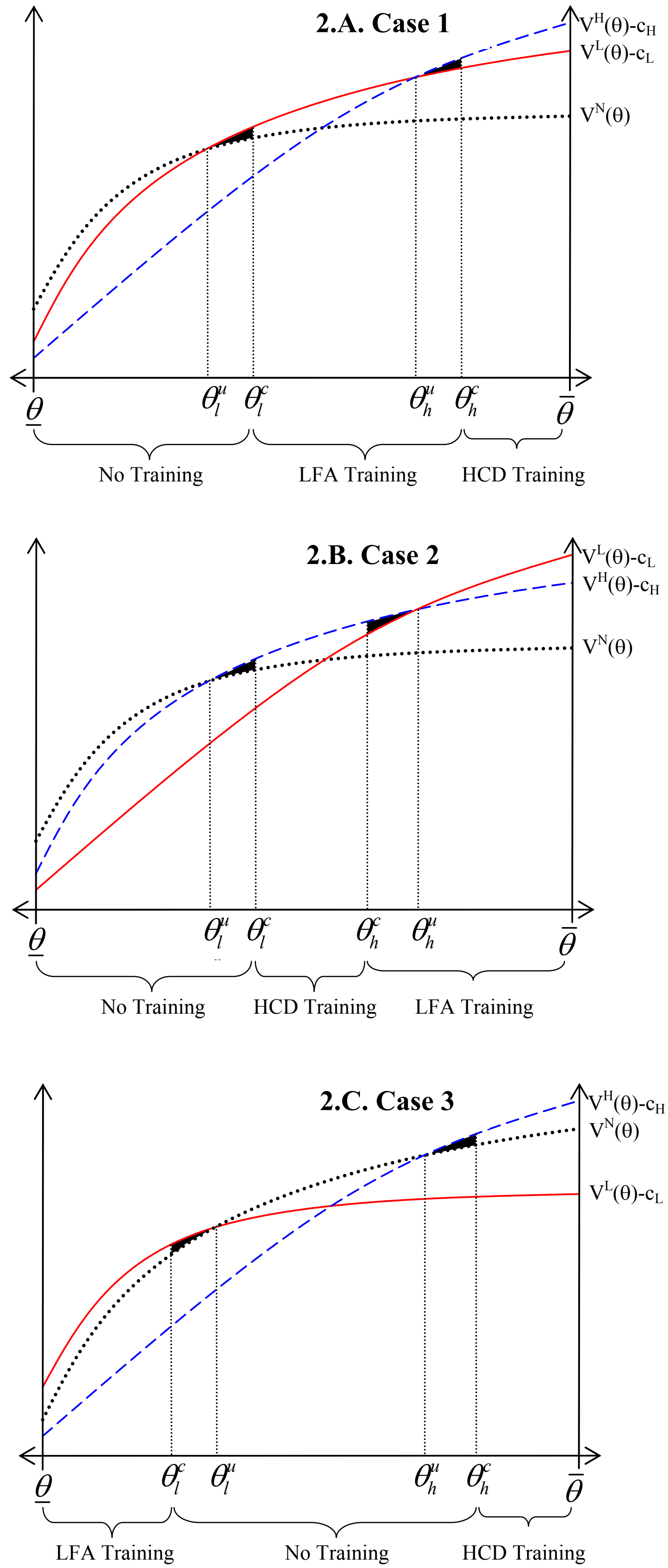
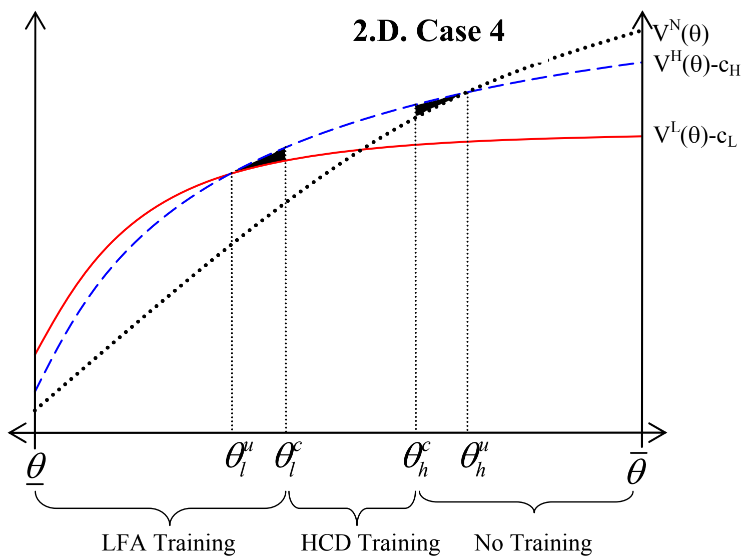
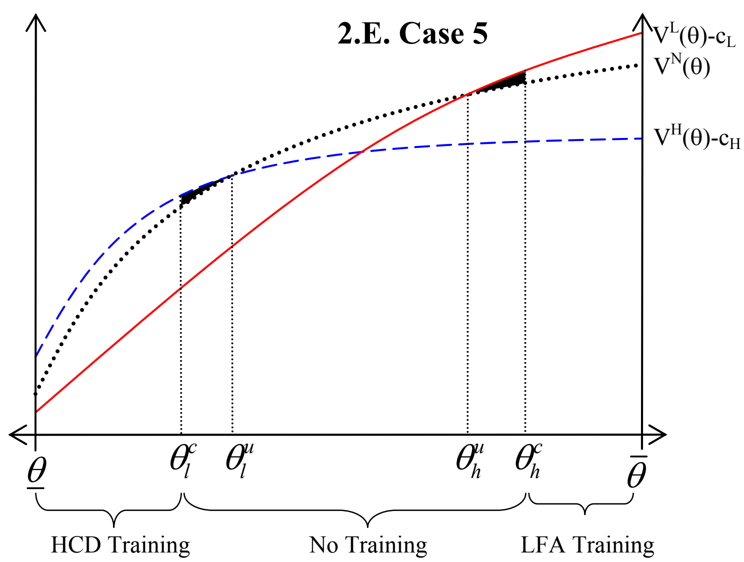


Figure 2. Interior Solution Cases for (P2) (continuation)





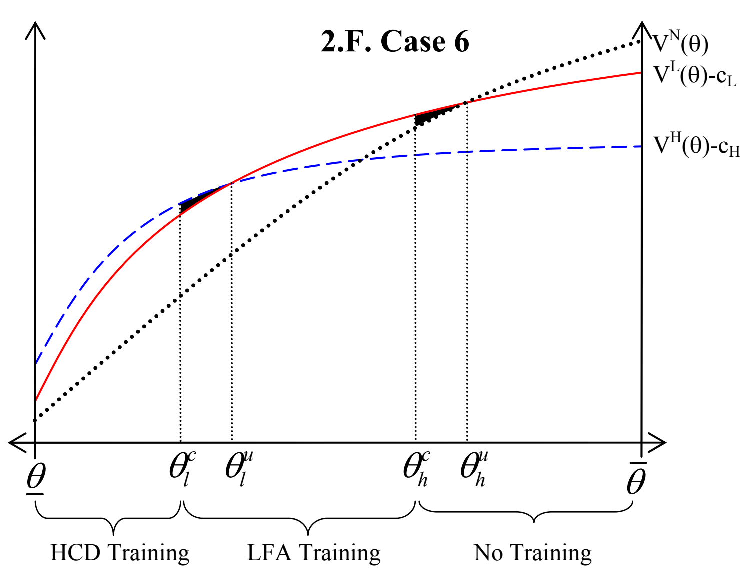


Figure 3
Density of Skills Distribution

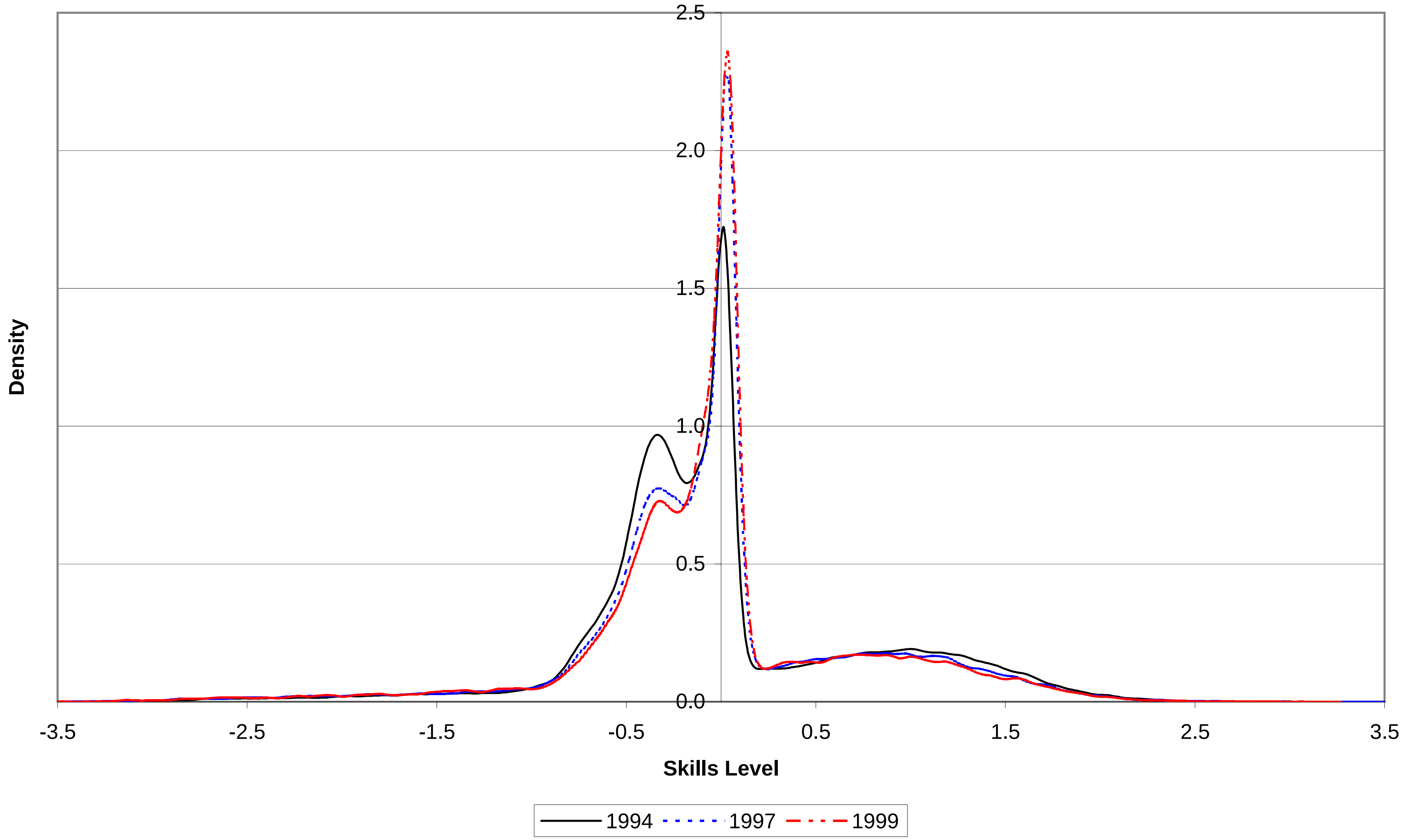


Figure 4
Mean Skills Level

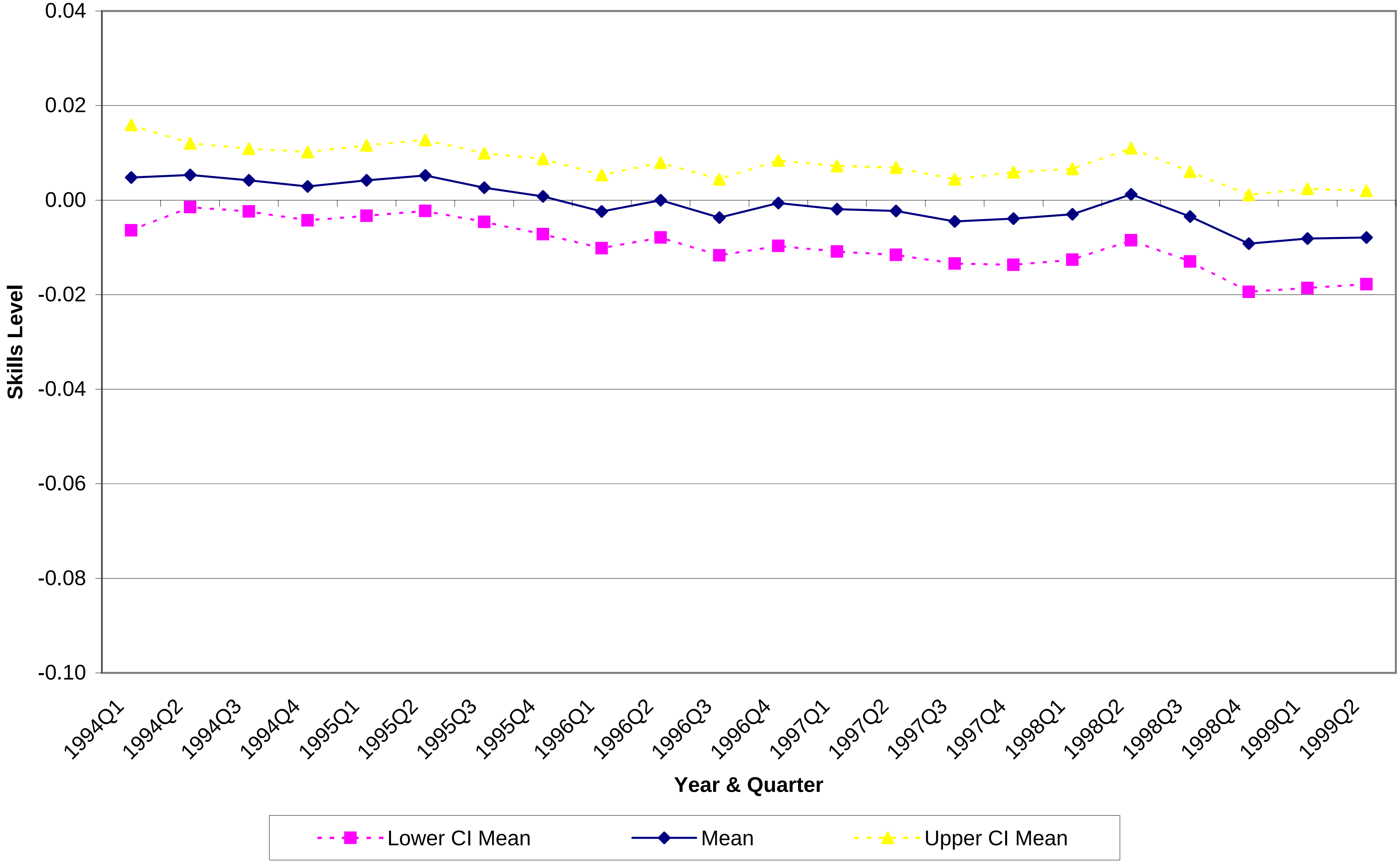


Figure 5
Median Skills Level

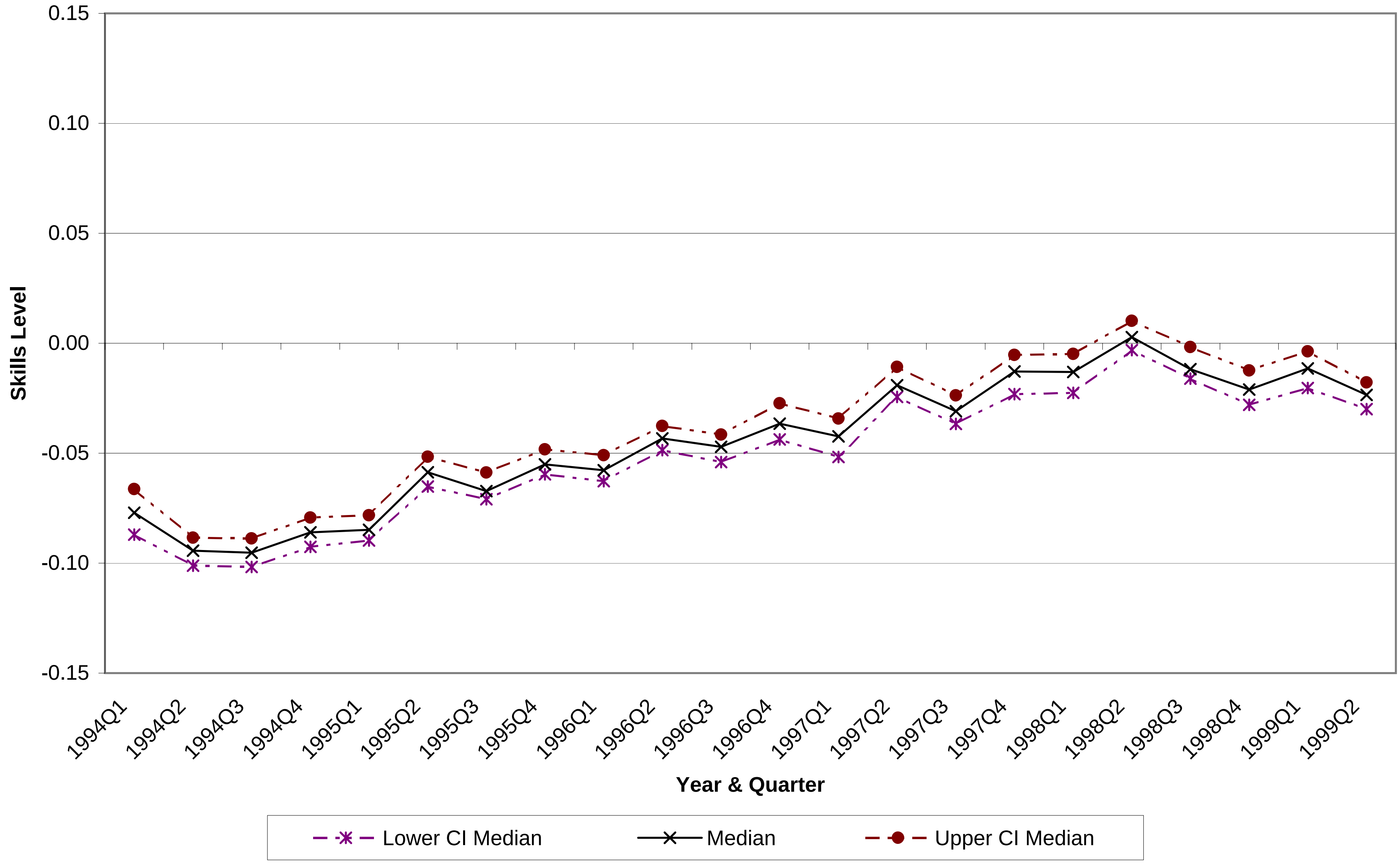


Figure 6
10th Percentile Skills Distribution

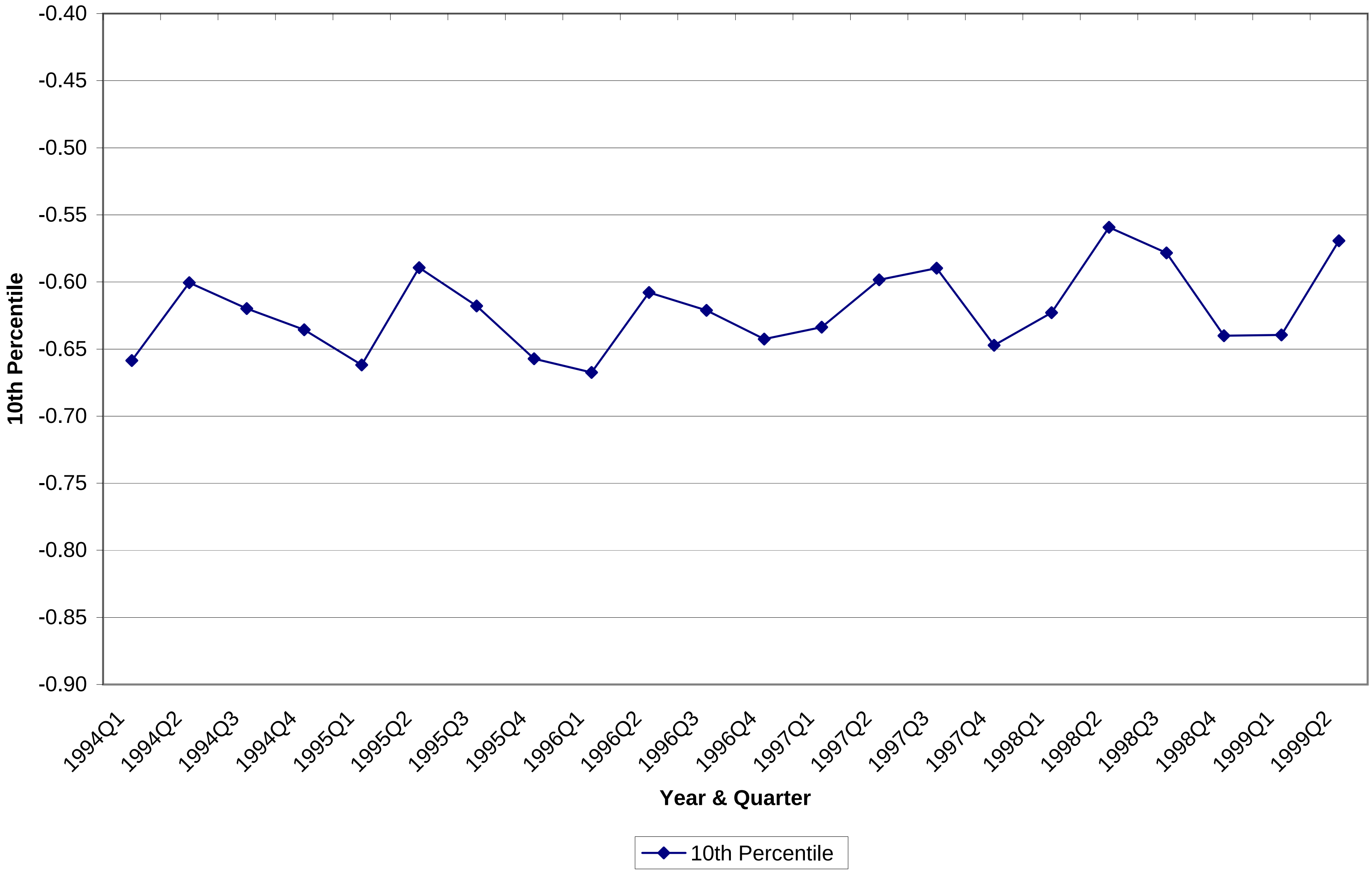


Figure 7
90th Percentile Skills Distribution

Appendix Table A.1. Proportions of adults on welfare trained - Annual averages

County	Proportion Not Trained						Proportion LFA Training						Proportion HCD Training					
	1994	1995	1996	1997	1998	1999*	1994	1995	1996	1997	1998	1999*	1994	1995	1996	1997	1998	1999*
Alameda	0.903	0.930	0.903	0.833	0.243	0.055	0.038	0.038	0.070	0.123	0.660	0.670	0.060	0.033	0.028	0.045	0.098	0.275
Butte	0.535	0.488	0.460	0.173	0.058	0.000	0.200	0.245	0.368	0.698	0.850	0.825	0.265	0.268	0.173	0.130	0.093	0.175
Contra Costa	0.843	0.775	0.778	0.618	0.358	0.000	0.093	0.135	0.185	0.345	0.618	1.000	0.065	0.090	0.038	0.038	0.025	0.000
Fresno	0.598	0.520	0.378	0.298	0.163	0.060	0.188	0.285	0.440	0.480	0.665	0.755	0.215	0.195	0.183	0.223	0.173	0.185
Imperial	0.493	0.503	0.465	0.353	0.100	0.000	0.293	0.310	0.328	0.370	0.475	0.535	0.215	0.188	0.208	0.278	0.425	0.465
Kern	0.540	0.578	0.523	0.450	0.313	0.015	0.308	0.263	0.360	0.443	0.568	0.795	0.153	0.160	0.118	0.108	0.120	0.190
Los Angeles	0.858	0.868	0.865	0.818	0.688	0.455	0.073	0.095	0.108	0.150	0.240	0.385	0.070	0.038	0.028	0.033	0.073	0.160
Madera	0.415	0.448	0.485	0.435	0.308	0.290	0.328	0.335	0.393	0.445	0.523	0.370	0.258	0.218	0.123	0.120	0.170	0.340
Merced	0.625	0.613	0.463	0.473	0.330	0.070	0.165	0.160	0.333	0.403	0.550	0.750	0.210	0.228	0.205	0.125	0.120	0.180
Monterey	0.810	0.815	0.788	0.728	0.668	0.600	0.075	0.098	0.175	0.233	0.250	0.310	0.115	0.088	0.038	0.040	0.083	0.090
Orange	0.838	0.828	0.788	0.668	0.428	0.280	0.085	0.095	0.158	0.268	0.495	0.610	0.078	0.078	0.055	0.065	0.078	0.110
Riverside	0.600	0.558	0.540	0.528	0.303	0.270	0.263	0.343	0.415	0.453	0.598	0.495	0.138	0.100	0.045	0.020	0.100	0.235
Sacramento	0.880	0.890	0.910	0.890	0.763	0.610	0.053	0.053	0.055	0.088	0.210	0.320	0.068	0.058	0.035	0.023	0.028	0.070
San Bernardino	0.870	0.813	0.803	0.755	0.355	0.050	0.083	0.143	0.170	0.223	0.575	0.815	0.048	0.045	0.028	0.023	0.070	0.135
San Diego	0.763	0.730	0.720	0.600	0.198	0.000	0.165	0.200	0.218	0.328	0.588	0.690	0.073	0.070	0.063	0.073	0.215	0.310
San Francisco	0.910	0.905	0.753	0.630	0.363	0.115	0.045	0.070	0.220	0.330	0.123	0.265	0.045	0.025	0.028	0.040	0.515	0.620
San Joaquin	0.763	0.725	0.688	0.608	0.670	0.405	0.098	0.135	0.198	0.308	0.275	0.500	0.140	0.140	0.115	0.085	0.055	0.095
Santa Barbara	0.775	0.780	0.753	0.675	0.478	0.165	0.118	0.128	0.200	0.263	0.443	0.660	0.108	0.093	0.048	0.063	0.080	0.175
Santa Clara	0.830	0.803	0.670	0.665	0.738	0.660	0.065	0.128	0.198	0.230	0.145	0.180	0.105	0.070	0.133	0.105	0.118	0.160
Shasta	0.713	0.680	0.663	0.633	0.538	0.100	0.143	0.168	0.200	0.278	0.395	0.670	0.145	0.153	0.138	0.090	0.068	0.230
Solano	0.785	0.773	0.778	0.745	0.400	0.220	0.100	0.143	0.163	0.235	0.543	0.700	0.115	0.085	0.060	0.020	0.058	0.080
Sonoma	0.685	0.723	0.710	0.625	0.263	0.110	0.170	0.175	0.213	0.295	0.433	0.510	0.145	0.103	0.078	0.080	0.305	0.380
Stanislaus	0.685	0.485	0.478	0.540	0.385	0.260	0.200	0.298	0.355	0.335	0.538	0.685	0.115	0.218	0.168	0.125	0.078	0.055
Tulare	0.763	0.703	0.578	0.150	0.000	0.000	0.125	0.198	0.313	0.485	0.763	0.935	0.113	0.100	0.110	0.365	0.238	0.065
Ventura	0.793	0.698	0.618	0.588	0.283	0.200	0.115	0.213	0.308	0.320	0.268	0.060	0.093	0.090	0.075	0.093	0.450	0.740

Note:
* First two quarters only.

Appendix Table A.2. New Entrants California 1994 Q1 - 1999 Q2

A. Reasons for not including Individuals in sample

	Individuals	Percentage
Sample Size	449,636	64.3%
Missing/Invalid Information	208,052	29.8%
Valid Info but not in 25 counties	41,050	5.9%
Total Number Entrants	698,738	100.0%

B. Comparison data quality 25 counties versus rest

	25 Counties	Rest
Individual kept in sample	70.2%	70.5%
Mainly SSN not valid	5.6%	3.6%
Mainly case type not valid	18.6%	21.0%
Mainly Individual >45	4.3%	3.3%
Mainly missing demographic vars	1.3%	1.7%

C. Entrants kept in sample 1994 Q1 - 1999 Q2 by county

County	Individuals	Percentage
Alameda	16,151	3.6%
Butte	4,267	0.9%
Contra Costa	9,643	2.1%
Fresno	17,787	4.0%
Imperial	4,449	1.0%
Kern	16,607	3.7%
Los Angeles	138,060	30.7%
Madera	2,685	0.6%
Merced	6,102	1.4%
Monterey	6,113	1.4%
Orange	25,452	5.7%
Riverside	24,408	5.4%
Sacramento	25,361	5.6%
San Bernardino	35,696	7.9%
San Diego	36,586	8.1%
San Francisco	7,134	1.6%
San Joaquin	11,881	2.6%
Santa Barbara	5,021	1.1%
Santa Clara	16,149	3.6%
Shasta	4,159	0.9%
Solano	5,337	1.2%
Sonoma	4,557	1.0%
Stanislaus	9,145	2.0%
Tulare	9,459	2.1%
Ventura	7,427	1.7%
Total	449,636	100.0%

Appendix Table A3. First stage IV for training expenditures (selected variables)

Variable	Log(Training Budget)
Log(Total Budget)	1.06***
	(0.10)
Proportion Votes Democratic Party	-0.03
	(0.12)
Proportion Adults Registered Democrat	-0.85
	(1.43)
Proportion Adults Registered Indep/Other	-0.46
	(2.04)
Proportion Adults Registered & Hispanic	-5.69***
	(1.47)
Proportion Adults Registered & Female	1.60
	(3.25)
Proportion Adults Registered & 18-34 yrs old	0.02
	(1.14)
Proportion Adults Registered & 55+ yrs old	-1.04
	(2.19)
Proportion Adults Registered to Vote	0.43
	(2.45)
Proportion Adults Voted in Election	0.20
	(0.17)
Observations	529
R-squared	0.99
Controls for LEC, Demographic, Skills Distribution	Yes
County Fixed Effects	Yes
Fiscal Year Fixed Effects	Yes
F test political vars=0 (p-value)	0.01

Note:
Standard errors between parentheses (adjusted by county/fiscal year clusters)
* significant at 10%; ** significant at 5%; *** significant at 1%