The Effects of Youth Employment: Evidence from New York City Lotteries[†]

Alexander Gelber
Goldman School of Public Policy, UC Berkeley, and NBER

Adam Isen
Office of Tax Analysis, U.S. Department of the Treasury

Judd B. Kessler
The Wharton School, University of Pennsylvania

Abstract

Programs to encourage labor market activity among youth, including public employment programs and wage subsidies like the Work Opportunity Tax Credit, can be supported by three broad rationales. They may: (1) provide contemporaneous income support to participants; (2) encourage work experience that improves future employment and/or educational outcomes of participants; and/or (3) keep participants "out of trouble." We study randomized lotteries for access to the New York City (NYC) Summer Youth Employment Program (SYEP), the largest summer youth employment program in the U.S., by merging SYEP administrative data on 294,580 lottery participants to IRS data on the universe of U.S. tax records; to New York State administrative incarceration data; and to NYC administrative cause of death data. In assessing the three rationales, we find that: (1) SYEP participation causes average earnings and the probability of employment to increase in the year of program participation, with modest contemporaneous crowdout of other earnings and employment; (2) SYEP participation causes a modest decrease in average earnings for three years following the program and has no impact on college enrollment; and (3) SYEP participation decreases the probability of incarceration and decreases the probability of mortality, which has important and potentially pivotal implications for analyzing the net benefits of the program.

[†] We thank the New York City Department of Youth and Community Development for sharing the data with us. We are particularly grateful to Alan Cheng for answering our questions. We thank the New York City Department of Health and Mental Hygiene for giving us access to their data. We thank Lawrence Katz (the Editor), David Autor, David Card, Hilary Hoynes, Robert Kaestner, Pat Kline, Dara Lee Luca, Janet McCubbin, Enrico Moretti, Steve Raphael, Jesse Rothstein, Chris Walters, Danny Yagan, and four anonymous referees for helpful suggestions, and we thank the Human Resources Center at The Wharton School and the UC Berkeley Institute for Research on Labor and Employment and Center on Governing and Investing for the Future for financial support. We thank Gita DeVaney, Harsha Mallajosyula, and Felix Owusu for outstanding research assistance. This paper does not necessarily reflect the views of the U.S. Treasury. All errors are our own.

1. Introduction

Many policies attempt to support individuals' labor market prospects, including public employment and subsidized employment programs. Youth unemployment in particular remains stubbornly high following the Great Recession both in the U.S. — where the unemployment rate for 16-24 year-olds was 12.2 percent as of this writing in 2015 — and throughout much of the world. In light of high youth unemployment, policy-makers have increasingly scrutinized youth employment programs. City programs across the U.S. provide youth with summer jobs — the fifty most populous cities in the country have all had summer youth employment programs in the last five years — and the federal Work Opportunity Tax Credit (WOTC) subsidizes employment of summer youth employees. While literature typically finds that non-summer-employment active labor market programs for youth have costs that outweigh their benefits, summer youth employment has "received relatively little attention from program evaluators" (Lalonde 2003, p. 532).

Programs to support summer youth employment are justified with various rationales. One rationale is that summer employment could provide income support to youth (and their families) through wages earned in the program. The website of the New York City (NYC) Department of Youth and Community and Development (DYCD), which runs the Summer Youth Employment Program (SYEP) that we analyze in this paper, states that SYEP aims to "provide supplemental income to aid low income families."[1] Similarly, economic stimulus efforts often aim to increase contemporaneous net earnings and employment. A second rationale is that summer work experience could improve future employment outcomes, by directly increasing human capital — the NYC DYCD also states that SYEP aims to "develop youth skills" — by encouraging youth to receive more schooling after participating in the program, or by acting as a signal to potential future employers.[2] A third rationale for such programs is that they could help to keep youth involved in socially productive activities or "out of trouble."[3] Keeping youth out of trouble during the summer could have immediate

[1] See http://usmayors.org/workforce/documents/2010-7-01USCOMWDCSYEPPresentation011910.pdf. Accessed May 16, 2014.
[2] See http://www.nyc.gov/html/dycd/html/resources/syep.shtml. Accessed May 16, 2014.
[3] See http://nycfuture.org/events/event/summit-on-the-future-of-workforce-development-in-new-york-city. Accessed May 16, 2014.

benefits through incapacitation or could place youth on a safer path that leads to decreased incarceration or mortality rates later in life.

We investigate the empirical support for these three rationales by analyzing the SYEP program in the years 2005 to 2008. During these years, SYEP provided summer jobs to NYC youth aged 14 to 21, paid by the NYC government at a total cost of $236 million.[4] Each year, SYEP received more applications than the number of SYEP jobs available and randomly allocated spots in the program by lottery. We compare the outcomes of individuals who participate in SYEP because they were randomly selected to receive a job to the outcomes of those randomly not selected. We link SYEP administrative data on these lottery winners and losers to Internal Revenue Service (IRS) administrative data on the universe of U.S. federal tax data; to New York State (NYS) Department of Corrections and Community Services (DOCCS) administrative data on individuals incarcerated in NYS; and to NYC Department of Health and Mental Hygiene (DOH) administrative data on causes of death in New York City. In the four years of lotteries we study, there were 294,580 SYEP applications subject to the lottery, of which 164,977 won the lottery and 129,603 lost the lottery.

This context provides a promising setting for studying a youth employment program. The large scale of the program, the random assignment, and the accurate data allow us to estimate precise causal effects on earnings, the employment rate, college enrollment, mortality, and incarceration up to a decade after program participation. Our sample sizes are at least an order of magnitude (and in many cases two orders of magnitude) larger than other randomized studies. The ability to look precisely at mortality, which other studies have not been able to observe, will prove particularly interesting since it has important implications for the magnitude of program benefits. NYC SYEP is also the largest summer youth employment program in the U.S. and therefore represents a central, recent case study of U.S. summer youth employment programs, and of youth employment programs more generally.

We find that SYEP participation increases earnings and employment in the year of the program. In a baseline specification, SYEP raises average earnings through the program by $1,085.34 in the year of program participation, lowers other earnings by a modest $208.87, and therefore raises net earnings by $876.26. Thus, crowdout of other earnings was

[4] Except where otherwise noted, all dollar amounts reported are in real 2013 dollars.

19.24 percent of the SYEP transfer in this year. We also estimate that, on net, SYEP raises the probability of having any job by 71 percentage points in the year of the participation, with a five percentage point decrease in the probability of having a non-SYEP job.

We do not find that youth employment has a positive effect on subsequent earnings or on college enrollment. In each of the three years following SYEP participation, SYEP participation causes a modest *decrease* in earnings of around $100 per year. Starting in the fourth year following SYEP participation, SYEP participation has an insignificant impact on earnings. The negative earnings effect in those three years is observed primarily among youth who are relatively older and have some work experience, and participation had an insignificant impact on subsequent earnings among WOTC-eligible individuals. We also find that SYEP has no impact on college enrollment, with an extremely precise 95-percent confidence interval that rules out a positive or negative effect greater than one one-hundredth of a year of college. It is notable that even for this young group with typically little prior job experience, an employment program did not provide a path to greater future earnings.

Over the year of SYEP participation and the subsequent four years, participation on net raises average earnings by $536.53. Thus, SYEP on net transfers to youth, though with significant crowdout (54.02 percent) of other earnings. Crowdout of other earnings is small relative to likely lifetime earnings but is substantial relative to the size of the program.

Consistent with keeping youth "out of trouble," SYEP participation decreases the probability of incarceration and decreases the probability of mortality. SYEP reduces the probability of incarceration by 0.10 percentage points, driven by a decrease among males. While this effect is small in absolute terms, it represents a substantial 10.36 percent reduction relative to the baseline. The SYEP-induced decrease in mortality, also driven by males, is 0.08 percentage points, again small in percentage point terms but a substantial 19.92 percent of the baseline. Cause of death data show that SYEP prevents death from external causes.

The point estimates imply that by October 2014, around 86 lives were saved by the SYEP program from 2005 to 2008. Under standard cost-benefit analysis calculations, this implies benefits of $773 million. Past literature that has typically found negative net benefits of active labor market programs for youth but has not examined the mortality outcome (e.g. Hollister, Kemper, and Maynard 1984, Couch 1992, Bloom et al. 1997, Cave et al. 1993, or Hendra et al. 2011; see Stanley, Katz, and Krueger 1998, Heckman, Lalonde, and Smith

1999, Lalonde 2003, and Card, Kluve, and Weber 2010).[5] Like most previous work on such programs, we find that the effects on future earnings cannot justify the program in a cost-benefit analysis; in fact, we find that SYEP modestly reduces participants' subsequent earnings. Adding a new twist to previous work, our mortality results show a very large new source of benefits that could well be pivotal to the cost-benefit analysis.

Amid the extensive literature on active labor market programs, the literature on summer youth employment contains only a few studies. Criminologist Sara Heller (2014) examines a randomized controlled trial (N=1,634) and finds that a summer youth employment program, in some cases in combination with cognitive behavioral therapy, greatly decreased violent crime arrests, had no significant impact on arrests for property, drug, or other types of crime, and had little impact on schooling — but her paper does not examine (1) longer-term impacts past 16 months after the program or (2) the impact on mortality, earnings, college enrollment, or incarceration.[6] Our finding of negative effects on earnings in the years subsequent to the program echoes some of the findings in a more recent literature about temporary employment programs (not specifically for youth), such as Autor and Houseman (2010), Card and Hyslop (2005), or many "Work First" programs (Bloom and Michalopoulos 2001).

The paper is structured as follows. Section 2 describes the policy environment. Section 3 describes our empirical specification. Section 4 describes the data we use. Section 5 discusses the first stage and the validity of the lottery. Section 6 discusses our results on earnings and employment. Section 7 presents results on college enrollment. Section 8 shows results on incarceration. Section 9 discusses results on mortality. Section 10 concludes.

2. Policy Environment

During the years we study (2005 to 2008), SYEP provided NYC youth aged 14 to 21 with paid summer employment for up to seven weeks in July and August.[7] Since 2005,

---

[5] Job Corps shows negative net benefits for the full sample, with positive net benefits only for the older-youth subgroup (Schochet, Burghardt, and McConnel 2008; Lee 2009). Studies of WOTC have focused on the take-up of the program (Hamersma 2003) or focused on those eligible for WOTC through long-term welfare receipt.
[6] See also Farkas et al. (1984), Crane and Ellwood (1984), Grossman and Sipe (1992), and McClanahan, Sipe, and Smith (2004).
[7] See, for example, the SYEP annual report from 2007 http://www.nyc.gov/html/dycd/downloads/pdf/syep_2007_annual_summary.pdf (accessed August 4, 2014).

DYCD has stored computerized records of applications, which were made available for this research. Because SYEP ran the program on its own, we are evaluating an existing government program (as opposed to a randomized experiment designed by researchers).

SYEP places participants in entry-level jobs and pays them the NYS minimum wage for working up to 25 hours per week during the summer.[8] In 2005 to 2008, the mean expenditure per SYEP participant per time participating in SYEP was $1,403 (including both wages paid to participants and administrative costs).

SYEP provides youth with various types of jobs, including jobs at summer camps, daycare centers, government agencies, hospitals, law firms, and museums. Nearly half of SYEP jobs are at summer camps or day care centers. In 2005 to 2008, 74.68 percent of the jobs were with non-profit, private sector firms; 10.95 percent were with for-profit, private sector firms; and 14.37 were with government entities. Thus, the program is typically closer to a "Work Experience" program, in which individuals are given temporary private sector jobs, than to "Public Sector Employment" program, in which individuals are given a government job (e.g. Heckman, Lalonde, and Smith 1999). The jobs that participants perform vary widely across employers but typically involve low-skill tasks. As an example of the most common jobs — those at summer camps of community organizations — one large, representative community center employer had five types of jobs available: camp counselor (who leads activities with children like song, dance, and physical activities); group leader (who leads counselors); support staff (who assist camp staff in daily activities like distributing lunch); clerical aide/office assistant; and janitor assistant/custodian (personal correspondence with DYCD, 3/17/2015).

All NYC youth who provide certain documentation are eligible to apply for SYEP. Applicants must show proof of identity using an official picture ID; proof of employment authorization; proof of age; proof of Social Security Number using a Social Security card; working papers for those under 18 (a Blue Card for those 14-15 and a Green Card for those 16-17); proof of citizenship/alien status; proof of address; and proof of family income. Males 18 and older must show proof of Selective Service registration.[9]

---

[8] In the years of our data the nominal state minimum wage rose from $6.00 per hour in 2005 to $6.75 per hour in 2006 to $7.15 per hour in 2007 and 2008. In 2014, it is $8.00 per hour. SYEP does not pay for overtime.
[9] See, for example, https://application.nycsyep.com/Images/SYEP_2014_Required%20Documents.pdf.

SYEP is administered by community-based organizations called "providers," which contract with DYCD to place SYEP participants into worksites and administer the program. Participants typically do not work directly for providers, but rather work for the employers to which providers match participants. In 2005 to 2008, the mean number of SYEP participants working for a given SYEP employer was 5.69. Over the summer, providers give participants around 17.5 hours of workshops on job readiness, career exploration, financial literacy and opportunities to continue education, or roughly 10 percent of the total hours in SYEP. During the years we study, this training component was decentralized across providers, was typically not considered a crucial component of the program, and generally was not costly for providers to deliver (personal correspondence with DYCD, 3/17/2015).

In a given year, applicants to SYEP apply through a specific SYEP provider. Individuals choose the provider to which they apply; applicants typically choose a provider located near their home. In a given year, an applicant applies to only one provider and is unable to apply to other providers at any point in that year. The application period is usually early-April to mid-May of the program year. Since there are more applicants than available slots in each year, the individuals who are allowed to participate in SYEP are selected by lottery. *Within* each provider in each year, there is a lottery to determine which individuals are selected for SYEP. Thus, winning the lottery is random *conditional* on applying to a given provider in a given year.

In each year, SYEP selected applicants through a series of lotteries. In an initial lottery, SYEP randomly selected winners and losers, where the number of winners was chosen to match the number of SYEP jobs available. However, not all of the individuals selected through this initial lottery participated in SYEP. Selected individuals may have chosen not to participate or failed to prove eligibility to participate. To fill the remaining slots, SYEP providers conducted subsequent lotteries. In each lottery, the number of winners was selected to match the number of remaining jobs at the SYEP provider, until the number of SYEP enrollees approximately matched the number of available jobs. We obtained data from SYEP on both the winners and losers of the initial SYEP lottery, and (separately) on the identities of those who won *any* of the lotteries in a given year and provider (as well as the identities of those who lost all lotteries in a given year and provider). For an applicant to a given SYEP provider, if a lottery occurred and s/he had not won a slot yet or had won a slot

previously but did not accept it, s/he was automatically entered into the subsequent lottery. Individuals were not able to withdraw their applications after the application deadline, nor were they able to enter subsequent lotteries if they had not applied to the provider by the deadline. Since selection of individuals was random in every lottery conditional on reaching that lottery, the dummy for whether an individual won any of the lotteries is exogenous. In our baseline specification, our instrument is a dummy for winning any of the lotteries.[10]

In any given year, individuals not selected in any of these lotteries were officially not able to participate in SYEP in that year, though they remained eligible to apply to SYEP in a subsequent year. Winning or losing the lottery in a given year, or participating in SYEP in a given year, does not affect the probability of winning or losing the lottery in a subsequent year, conditional on applying in the subsequent year. The opportunities to participate in comparable government programs are small relative to the size of SYEP.[11]

Providers make the assignments of participants to employers, and to particular jobs within employers, based on two inputs. Applicants specify their skills and industry interests on their applications, and employers inform providers of restrictions on the type of participants they can hire (*e.g.* an employer might require high school graduates for certain jobs). The particular method for matching participants to jobs based on these two inputs varies across providers. Once a provider matched a participant to a job within an employer, the firm occasionally chose to re-allocate the participant to a different job during the course of the summer.

3. Data

DYCD data

The DYCD data on SYEP contain a number of key pieces of information that we use, including: whether an individual won or lost any of the SYEP lotteries (including the subsequent lotteries); whether the individual participated in SYEP; which provider an individual applied to; the year the lottery was conducted; self-reported information on variables including gender, date of birth, race, and name; and Social Security Number (SSN).

[10] As shorthand, we sometimes refer to "winning (losing) any of the lotteries at a given provider in a given year" as "winning (losing) the lottery." In the Appendix we show that the results are similar when our instrument is a dummy for winning the initial lottery (which is a slightly less powerful instrument).

[11] http://www.nyc.gov/html/dycd/downloads/pdf/Summer_Youth_Alternatives2014.pdf (accessed 8/4/2014).

The data include information on all SYEP applicants, regardless of whether they enrolled in SYEP or not. For SYEP participants, the data additionally include the industry the individual worked in through SYEP (in industry categories created by SYEP).

IRS data

We merge the SYEP administrative data to IRS administrative data using SSN, which matches 99.6 percent of the SYEP applicants to the IRS data. It is not surprising that we obtain a very high match rate, as individuals were required to list their Social Security number and show their Social Security card (as well as the voluminous additional documentation listed above) to be eligible for SYEP. To include additional individuals who may have an incorrect SSN listed but have other information correct, we match the remaining SYEP data to IRS data when name, gender, day of birth, month of birth, year of birth, and first or last four digits of the SSN all match. This allows us to match an additional 0.2 percentage point of the SYEP data to the IRS data, for a total match rate of 99.8 percent. The results are robust to other matching procedures.

The IRS data contain a wide variety of information including: name; date of birth; age; gender; the identity of their family members; the Employer Identification Numbers (EINs) of their employer(s); the North American Industrial Classification System (NAICS) industry code of their employer; each individual's day, month, and year of death (if any); whether an individual's employer is a non-profit; and whether the individual is enrolled in college. Our measure of an individual's annual earnings comes from W-2s, mandatory information returns filed with the IRS by employers for each employee for whom the firm withholds taxes and/or to whom remuneration exceeds a modest threshold.[12] Thus, we have data on W-2 earnings regardless of whether an employee files taxes. Having "any job" is defined as earning a positive amount, again as reported on the W-2 form. Mortality is observed for the full population, and college enrollment is observed on a mandatory information return, and therefore neither depends on filing status. We use data on each of these variables in each year from 2004 to 2012 (inclusive). We winsorize earnings at

[12] This measure does not include self-employment income, as reported on the 1099-MISC or 1040 Schedule C. We find negligible impacts of SYEP on these measures of self-employment income.

$100,000 (*e.g.* Chetty *et al.* 2011). Like most administrative datasets, the data lack information about the hourly wage, hours worked, or underground earnings.

SYEP participants receive a W-2 from the NYC government (rather than the employer they worked for through SYEP). Earnings from the NYC government is an extremely good measure of their earnings through SYEP. In principle, this measure could differ from their earnings through SYEP if the individual also held another job with the NYC city government, but mean NYC government earnings among non-participants is very small.

Incarceration Data

We also collected data from the NYS DOCCS on individuals incarcerated in a NYS prison in years up to and including 2013. Everyone who has been confined in NYS prison is listed in the database, except those who were 18 or under at the time the offense was committed, those who have had their convictions reversed by a court, and certain offenders who are covered by a special provision for relatively minor crimes. The exclusion of youthful offenders is a particularly important data limitation in our context, because most SYEP applicants are 18 or younger at the time of SYEP application and therefore incarceration episodes that occurred due to crimes around the time of SYEP participation will be excluded from the data (even once they have reached age 19). We also do not observe those jailed in a local jail such as Riker's Island. In total, we observe 466,062 unique incarceration episodes in the DOCCS data. Since the DOCCS data do not include SSN, we match information from the DOCCS data to the SYEP administrative data on when first name, last name, day of birth, month of birth, and year of birth all match. 0.95 percent of SYEP applications, and 1.01 percent of SYEP applicants — a total of 2,004 SYEP applicants — match to the DOCCS data. Among those incarcerated, 93.79 percent were incarcerated once.

Cause of Death Data

Our main estimates on mortality use the IRS mortality data previously described, which cover the full U.S. population through October 2014. To further investigate the observed effect on mortality, we matched the SYEP data to NYC Department of Health and Mental Hygiene (DOH) administrative data on the cause of death for individuals who died from known causes in NYC from 2005 to 2013 (the most recent year of data currently

available). Just as with the IRS data, we match SYEP earnings records to DOH data on the basis of SSN, first name, last name, and month, day, and year of birth. We match 891 unique DOH mortality episodes (occurring from 2005 to 2013 in NYC) to the SYEP data.

Data setup

In the discussion that follows, we call "Year 0" the year an individual applies to a SYEP lottery. In Year 0, an individual participates in the SYEP program (if they win the lottery and take the SYEP job), or alternatively they do not participate. Year 1 refers to the following calendar year, Year -1 refers to the year before Year 0, and so on.[13] In the years we examine, SYEP gave "special slots" for disabled youth that were not selected by lottery. We drop these applicants from our sample. We also delete observations in which the same SSN is associated with multiple applications in a given year, thus deleting approximately 1,000 observations per calendar year in Years -1 to 4 (and fewer in subsequent years). The number of remaining observations in each year from Year -1 to Year 4 is 294,580, corresponding to 198,745 individuals. The number of observations in each year of data is greater than the number of individuals because some individuals apply to SYEP in more than one year. 114,013 individuals participated in SYEP at some point.

Because individuals can apply to SYEP in more than one year, our setup of the data follows the parallel setting in Cellini, Ferreira, and Rothstein (2010), in which treatment in a given year can affect the probability of treatment in a following year. Following their method, we stack multiple panels of data. In each panel, Year 0 is defined as the year an individual participates in a lottery. Thus, an individual appears in multiple panels if she applied to SYEP multiple times.

In any given year over the lottery years we study, around four percent of the eligible population in NYC participated in SYEP. Since we have complete IRS data until 2012, we observe everyone until at least Year 4 (as the last lottery we observe is in 2008).

Summary statistics

[13] For example, for individuals in the 2005 lottery, Year -1 refers to the Year 2004; Year 0 refers to the Year 2005; Year 1 refers to 2006; and so on.

The first column of data in Table 1 shows summary statistics for the full sample. Given applicants' young ages, it is not surprising that mean total earnings over Years 0 to 4 are quite low compared to the general population — only $3,555.29. Mean NYC government earnings over this period are $218.39.[14] 63 percent are employed in any job, and 50 percent have any non-NYC government job. 23 percent are enrolled in college in a given year. Median total earnings (shown in Appendix Table 9) rises from $939.18 in Year 0 to $2,473.01 in Year 4. Turning to the demographics of SYEP applicants, on average they come from disadvantaged family backgrounds and are disproportionately minorities. Mean family income is low ($39,521.56 in Year -1).[15] 48 percent of SYEP applicants are black, far greater than the share of NYC residents who are black, whereas 13 percent of SYEP applicants are white, far lower than the share of NYC residents who are white. Just under one-half (45 percent) are male. The mean age is 16.50. The vast majority, 93 percent, are U.S. citizens.

Appendix Table 1 shows the breakdown of SYEP jobs by industry, as reported by SYEP.[16] SYEP reports that much of the sample works at a day care or day camp (36.99 percent) or at a camp outside of New York City (10.59 percent). While the SYEP industry classification is not based on the North American Industrial Classification System (NAICS), we use the descriptions provided by SYEP to develop a set of 2-digit NAICS codes that corresponds roughly to the industries described by SYEP (the crosswalk is shown in the Appendix). To classify SYEP jobs as for-profit, non-profit, or government, we use data reported by SYEP. For jobs not through SYEP, we classify employed individuals as working at a non-profit if their employer files a form 990; as working in the government if their NAICS code is 92; and as working for a for-profit otherwise.

4. Empirical Strategy

[14] Mean NYC government earnings in Years 0 to 4 are $218.39—lower than its mean in Year -1 ($256.96). The mean in Year -1 is higher because some of those who apply to SYEP in Year 0 participated in previous years and the mean in Years 0 to 4 is pulled down by Year 0 applicants reaching ages with lower participation rates.
[15] The 2011 American Community Survey reports that mean U.S. household income is $69,821.
[16] DYCD did not receive records of the EINs of the firms at which SYEP participants worked. Thus, we are limited to using the industry breakdown provided by SYEP.

Our empirical strategy exploits the random assignment of SYEP access through the lotteries. Since some of those selected for SYEP did not enroll, we use winning the SYEP lottery to instrument for participation. A basic two-stage least squares specification is:

$$P_{ij0} = \alpha_1 W_{ij0} + X_j\alpha + u_{ij0} \qquad (1)$$

$$E_{ijt} = \beta_1 P_{ij0} + X_j\beta + v_{ijt} \qquad (2)$$

Here E_{ijt} is a year-t outcome (such as the level of earnings in year t) of individual i that participated in SYEP provider lottery j. W_{ij0} is a dummy for whether the individual won the SYEP lottery or not in Year 0. P_{ij0} is a dummy for whether the individual participated in SYEP in Year 0. Because individuals applied to providers and the lotteries were run at the provider level in each year of the lottery, we control for a vector X_j of dummies for each provider in each year of the lottery. In some specifications, we control for additional covariates. u_{ij0} and v_{ijt} are error terms. We cluster our errors by SYEP provider, which we view as a conservative choice. There are 62 providers in our data. Clustering our standard errors at the individual level instead leads to nearly identical standard errors (Appendix Table 6). We interpret our coefficient β_1 as a local average treatment effect of SYEP among the compliers (*i.e.* those induced to participate in SYEP by winning the lottery).[17]

We typically investigate the results separately in different years, running our specification for each year t of outcomes separately. In some cases, we examine the results across multiple years (e.g. examining the effect of SYEP on total earnings in Years 0-4). In this case, we sum earnings across all of the years examined (e.g. summing earnings across Years 0-4) and run our specification with this summed earnings variable as the outcome.

It is possible that SYEP participation in Year 0 could affect the probability of applying to SYEP or the probability of accepting the SYEP job conditional on winning the lottery — and thus could affect SYEP participation — in subsequent years. In this case, part of our estimate of the effect of Year 0 SYEP participation on subsequent earnings (defined as earnings in calendar years following the calendar year of SYEP participation) could be mediated through the impact of SYEP on future SYEP participation. In the terminology of

[17] Because lottery losers are officially ineligible to participate, this also should represent the average treatment effect on the treated. However, in very rare cases (1.7 percent), lottery losers participated in SYEP, for example because after running all of the lotteries, providers still had remaining slots available and allocated remaining slots in the program to lottery losers.

Cellini, Ferreira, and Rothstein (2010), the specification in (1)-(2) is a "static" specification, in which we estimate the total effect of Year 0 SYEP participation on earnings in a given year, *including* effects that are mediated through the channel of the effect of SYEP participation in Year 0 on SYEP participation in subsequent years. In the Appendix, we also estimate the effect of SYEP participation on earnings using the "dynamic" design of Cellini, Ferreira, and Rothstein. In our context, this dynamic estimator effectively yields the effect of SYEP participation in Year 0 on earnings in any given year, while removing the effect that operates through the channel of the effect of Year 0 SYEP participation on subsequent SYEP participation. These two objects of study reflect different conceptual experiments of interest. However, since SYEP participation in Year 0 only slightly affects the probability of SYEP participation in subsequent years (i.e. Years 1-4), the static and dynamic estimates prove to be similar.

In some cases, we investigate a binary dependent variable, like a dummy for whether an individual has a job. In an instrumental variables model with a binary endogenous variable and a binary outcome, models such as a two-stage probit are in general inconsistent, and we run a linear probability model instead (Angrist 2001).[18] When we examine a binary variable pooled across years (e.g. probability of having a job, Years 0-4), we define the variable as the probability that the outcome occurs at any point during those years (in the example, the outcome is the probability that an individual has a job at any point in Years 0-4).

5. Preliminary Empirical Results

Validity of Randomization

Table 1 demonstrates the validity of the randomized design by comparing the characteristics of SYEP lottery winners and losers. We run a "reduced form" OLS regression of characteristics of SYEP applicants on a dummy for winning the lottery and provider-by-year fixed effects. We examine outcomes in the year prior to applying to SYEP and a number of demographic variables. Consistent with the validity of the randomization, none of these variables is significantly related to treatment status. Though not tabulated, we also find insignificant estimates in every other year prior to SYEP enrollment. The probability that

[18] The coefficients in the linear first stage and reduced form regressions are typically nearly identical to the marginal effects in the probit reduced form and first stage, and also to those in a bivariate probit.

SYEP applicants match to the IRS data is also balanced. A joint test of significance of all coefficients on treatment across all variables shows p=0.59.

First stage

Table 2 shows the first stage — the effect of winning the SYEP lottery in Year 0 on SYEP participation in Year 0 — as well as the effect of winning the SYEP lottery in Year 0 on SYEP participation in subsequent years. For Year 0 participation, the coefficient on the dummy for winning the SYEP lottery is 0.73, and the F-statistic is 4,188.68. As the take-up rate is 73 percent, the treatment-on-the-treated (TOT) estimates of the effect of Year 0 participation will generally be 1.37 (=1/0.73) times as large as the intent-to-treat (ITT) estimates.[19] SYEP participation in Year 0 affects the probability of SYEP participation in Years 1 to 4 separately, though these effects are very small (3 percentage points or less). Specification (1)-(2) restricts the first stage to be the same across all providers in all years of the lottery; Appendix Table 20 shows that our key results are extremely similar when we allow the first stage to be different across providers or provider-years.

Comparison of compliers and never-takers

Appendix Table 2 compares pre-determined characteristics between lottery winners who participated (*i.e.* compliers), and lottery losers who did not participate (*i.e.* never-takers). Compliers have lower average earnings and a lower probability of being enrolled in Year -1, are younger, are more likely to be black, and are more likely to be U.S. citizens.

6. Effects on Earnings

Main estimates of effects on earnings and probability of having a job

Table 3 shows our main estimates of the effect of SYEP participation on earnings and the probability of having a job. The point estimate of the effect of SYEP participation on total earnings in Year 0 is $876.26 ($p$<0.01), as SYEP participation on average leads to a substantial increase in earnings in the year of SYEP participation. This represents a near doubling of earnings relative to the control group mean. SYEP participation causes Year 0

[19] Our empirical strategy could also be used to examine the effect of employment in Year 0 (through SYEP or other employers) on earnings. In this case, we would scale up the linear estimates by a factor of 1.43.

earnings from the NYC government to increase by an average of $1,085.34 ($p$<0.01). SYEP participation reduces Year 0 non-NYC government earnings by $208.87 ($p$<0.01), or 19.24 percent of the increase in NYC government earnings.[20] SYEP participation raises the probability of having a job in Year 0 (including in both SYEP and non-SYEP jobs) by 71 percentage points. SYEP participation lowers the probability of having a non-NYC government job by 5 percentage points in Year 0, indicating modest crowdout. It is notable that crowdout of non-NYC government earnings was 19.24 percent, whereas crowdout of other jobs was only 5 percentage points. Evidently, conditional on having a job in Year 0, SYEP jobs tend to be lower-earning jobs.[21]

In each year from Year 1 to Year 3, SYEP participation in Year 0 lowers total earnings modestly, by around $100 ($p$<0.05) in each of the years separately. Relative to mean earnings in the control group each year, these negative effects on earnings represent earnings decreases of 4.44 percent, 2.90 percent, and 2.48 percent in Years 1, 2, and 3, respectively. From Years 1 to 3, SYEP participation raises NYC government earnings slightly (p<0.01); lowers non-NYC government earnings modestly (p<0.01); and lowers the probability of having a non-SYEP job slightly (p<0.01 in Years 1 and 2, and p<0.05 in Year 3). SYEP slightly raises the probability of having any job in Year 1. This combination of results again suggests that SYEP leads individuals to earn less conditional on having a job. The effect on total earnings in Year 4 turns insignificant, with a small confidence interval. These results are consistent with the Card, Kluve, and Weber (2010) meta-analysis findings that programs for youth, and programs involving subsidized jobs, often do not to have positive impacts on labor market outcomes.

The effect of SYEP on total earnings in Years 0 through 4 is positive and substantial ($536.53). The effect on total earnings is less than half of the average total of SYEP transfers over this period ($1167.30); the average decrease in other earnings is 54.02 percent of average SYEP earnings. There is also a positive effect of 9 percentage points on the probability of having any job during these years. Finally, the impact on total earnings in

[20] Some of this decrease in other earnings could have occurred in Year 0 *after* the summer of Year 0.
[21] The coefficient on the SYEP participation dummy when total number of jobs is the dependent variable is similar to the effect of SYEP participation on the probability of having a job, suggesting that the effect on holding multiple jobs is not responsible for this reduction in earnings conditional on having a job. We do not have data on hours worked to determine how hours compare in SYEP jobs and other jobs.

Years 1 through 4 is negative and substantial, while the impact on the probability of having a job during this period is small and positive.

Appendix Table 3 shows the "intent-to-treat" estimates. The coefficients are 73 percent as large as the IV estimates in Table 3. When we remove the provider-year dummies and therefore do not rely on randomized variation (but control for all of the 13 pre-determined covariates listed in Appendix Table 4), the estimates differ dramatically: SYEP participation is estimated to *raise* total future earnings (*e.g.* over Years 1-4, SYEP is estimated to raise total earnings by $593.56, $p<0.01$).

While our main specification examined Years 0-4 (to hold the sample size constant across years, and because the estimates turn insignificant beginning in Year 4), Appendix Table 4 shows the estimates for Years 5, 6, and 7. As we might expect from random chance, one estimate is significant, though not robust: the estimate for earnings in Year 7 is positive and significant without controls, but it becomes insignificant when we add the controls. Over Years 0 to 7, the results are similar to those over Years 0 to 4.

Appendix 1 discusses a wide variety of variations on these basic results, including: adding controls to the regressions; using the initial SYEP lottery as the instrument; including only individuals who match according to SSN; clustering at the individual level; the dynamic specification of Cellini, Ferreira, and Rothstein; investigating the effect of SYEP separately for those who had or had not previously participated in SYEP; using a SYEP lottery win as an instrument for the total number of times participating in SYEP; and estimating the effect on other family members' earnings. Appendix Tables 5 through 8 show that we continue to find comparable results throughout these alternative specifications. Appendix Table 9 shows that winning the SYEP lottery raises median earnings in Year 0 and has generally positive effects on median earnings in subsequent years but negative effects in higher quantiles, the latter of which drives the negative earnings results in Years 1 to 3.

Appendix Table 10 shows that SYEP has a more significant and negative effect on subsequent earnings among those ineligible for WOTC than among those eligible (i.e. ages 16-17 living in an Empowerment Zone (EZ)); among whites than among other race groups; among older SYEP participants than younger; and among those who worked in Year -1 than

among those who did not.[22] We find an insignificant effect of SYEP on total earnings in the 2005-6 lotteries but a substantial negative and significant effect in the 2007-8 lotteries.

Effects on type of job

We investigate the effect of SYEP on earnings in different industries. As an illustrative exercise, we classify industries into two clusters: those in which the 2-digit industry represents a greater percentage of total jobs among SYEP-provided jobs than among jobs held by the control group (Cluster 1), and industries in which the opposite is true (Cluster 2). Appendix Table 1 lists the industries in each Cluster. Table 4 shows that SYEP participation leads to an increase in Cluster 1 earnings and employment both in Year 0 and in subsequent years.[23] Table 4 also shows that SYEP strongly raises earnings in non-profit firms in Year 0 and continues to raise these earnings modestly through Year 4 (with similar results for the probability of having a job). Earnings in for-profit firms are lowered by SYEP by around $100 per year in Years 0, 1, 2, and 3. Meanwhile, SYEP increases earnings in government jobs in Year 0 but modestly reduces government earnings in Years 3 and 4.

Interpreting the Earnings Results

While the negative effects on subsequent earnings are small relative to likely lifetime earnings, it is worth considering the reasons behind the arguably surprising result that SYEP participation decreases earnings among a young group with little prior work experience, even during the Great Recession. Our randomized design is well suited to determine the program's causal impacts, but less equipped to determine the mechanisms that mediate these impacts. Thus, we can say only whether the predictions of our hypotheses are consistent with the data.

SYEP could crowd out jobs that could have led to greater future earnings.[24] As we discuss in Appendix 2, we find that groups that experienced greater Year 0 crowdout also experienced greater decreases in subsequent total earnings, as we would expect if crowdout

[22] An EZ is an area with particularly high poverty and/or emigration.
[23] When we perform these regressions using the dynamic estimator of Cellini, Ferreira, and Rothstein (2010), we obtain very similar results, suggesting that the effect is not driven by SYEP participants reapplying to SYEP but instead by some stickiness in job choice (see the discussion in Appendix 2).
[24] Our results find only 19.24 percent earnings crowdout in Year 0, which may limit the potential quantitative importance of this explanation. In principle, however, it is possible that SYEP participation in Year 0 could also negatively affect future earnings relative to the counterfactual of having no job in the formal sector in Year 0.

of other experiences in Year 0 leads to decreases in subsequent earnings. Further, the subgroup analysis found more negative impacts for groups that were more likely to otherwise be working in Year 0 — *i.e.* older individuals and those with a job in Year -1.[25] Relatedly, SYEP decreases the probability that an individual continues working for a past employer (Appendix Table 11), raising the possibility that SYEP harms a participant's career development with an existing employer.

Appendix Table 12 shows the interaction between winning the SYEP lottery and the fraction of jobs in the SYEP provider in Cluster 1. The regressions suggest that a Cluster 1 (Cluster 2) job placement increases (decreases) earnings both during and after SYEP, further suggesting that the effect of SYEP on Year 0 job type is a culprit for the negative effect on subsequent earnings. However, heterogeneity in the effects across providers could instead be driven by other factors that happen to be correlated with the types of jobs in each provider.[26]

Appendix 2 discusses other potential explanations including income effects, time-inseparability of leisure, signaling, changes in the labor supply curve, and peer effects; the evidence on these mechanisms is mixed, though we cannot rule them out. The next section shows that the decrease in subsequent earnings is not driven by college enrollment.

7. Effects on College Enrollment

In principle, SYEP could affect schooling decisions. Schooling is an investment that could lead individuals to decrease earnings in the years immediately after SYEP participation, as individuals focus on academics or enroll in college, but raise earnings in the slightly-more-distant future. Table 5 investigates the effect of SYEP participation on college enrollment.[27] Table 5 reports results with the full sample for consistency with our other estimates, although the results are extremely similar when we limit the sample only to observations when individuals are 18 and over (as those under 18 rarely attend college). We find no significant impact throughout, with very small standard errors. When we estimate the effect of SYEP participation on total years enrolled in college in Years 0-4, the point estimate is -0.001, and the confidence interval rules out an increase or decrease in total years of

[25] These samples differ on average along many characteristics, so this evidence is merely suggestive.
[26] When we regress earnings in Years 1-4 on provider dummies interacted with the SYEP lottery win dummy, a joint F-test of equality of the coefficients across providers is just shy of significant (p=0.11).
[27] Our data lack a measure of whether individuals graduated from college.

college greater than *one-hundredth of a year*. These estimates are nearly identical when we control for covariates (Appendix Table 13) or with the dynamic specification.

If SYEP had a positive impact on high school attendance, this could reduce individuals' earnings while they are of high school age. However, a range of evidence fails to support this hypothesis. Using SYEP data from 2007, Leos-Urbel (2012) found that winning the SYEP lottery slightly decreased the probability that an individual attended high school the following school year, though this effect was significant only at the 10 percent level.[28] While we do not have data on high school attendance, we can indirectly investigate whether an effect on high school could drive our negative earnings results. Appendix Table 14 shows that among those older than 18, who are too old to have still been in high school after the summer of SYEP, SYEP decreased subsequent earnings much more than in the full population. Furthermore, if there were a significant positive impact on high school attendance or completion, then we might expect (1) a positive impact of SYEP on earnings several years later; (2) a larger negative impact on near-term earnings in the younger group than the older group; and an eventual positive impact on the probability of (3) college enrollment and (4) having a job. None of these predictions is observed in the data (in fact, (1) and (2) are the opposite of what we observe in the data).

8. Effects on Incarceration

Keeping kids "out of trouble" during the summer could lead them away from crime and reduce the probability of incarceration. In Table 6, the dependent variable is a dummy for whether an individual appears in the DOCCS incarceration database. To parallel our main specification for employment, we estimate a linear probability two-stage least squares model.

In the full population (Row A), we find that SYEP reduces the probability of incarceration by 0.10 percentage points.[29] This is a 10.36 percent reduction relative to the baseline incarceration rate of 0.95 percent. In combination with the number of SYEP

[28] That paper's main focus is on the correlation between SYEP participation and log days attending school *conditional* on attending school (and finds that participation is associated with a very small, 1 percent increase in days attending school conditional on attending school). However, it is difficult to interpret this correlation as the causal effect of SYEP participation on days attended because the sample attending school is selected (due to the negative effect of SYEP participation on the probability of high school attendance).

[29] If the treatment observations had better data quality than the control observations, this would bias us toward estimating a *positive* effect of SYEP on incarceration – the opposite of our finding.

participants, this implies that 112 fewer people were incarcerated by 2013 as a result of SYEP participation between 2005 and 2008. This result is notable in light of literature reviews that conclude that "work doesn't work" in reducing crime (Bushway and Apel 2012; Cook *et al.* 2014). Recall that only individuals 19 and older when they commit a crime that leads to incarceration in a NYS prison are included in our DOCCS incarceration data. In the full sample, the estimates thus incorporate possible effects on incarceration for future crimes (for both those under and over 19 at the time of SYEP participation) and on incarceration for crimes committed simultaneously with SYEP participation (only for those 19 and older at the time of SYEP participation). In Appendix Table 15 we show that the results are very similar when controlling for covariates, when the dependent variable is number of times incarcerated, and with a probit.

We find important differences in the incarceration effect across subgroups. While we do not observe the timing of the crime committed, we find that SYEP causes a dramatic reduction in the incarceration rate among those who are 19 or older in the summer they participate in SYEP, among whom both incapacitation effects and effects on future behavior are possible. The reduction in the incarceration rate due to SYEP for this group is 0.46 percentage points ($p<0.05$) — a very large, 54 percent reduction relative to the baseline rate of 0.85 percentage points. In the group 18 and under when they participate in the program, the estimated effect is smaller and not quite statistically significant at the 10 percent level ($p=0.12$). The point estimates also suggest that SYEP reduces incarceration more among males than among females; more among those without prior work experience than those with prior experience; more among blacks and whites (particularly blacks) than among Latinos and other races; and more among those outside EZ's than those in them — though the treatment effects are only significantly different across subgroups in comparing males and females. The effects are significantly different across providers ($p<0.01$). Appendix Table 16 shows the results for subgroups within in the 19-and-older group.

9. Effects on Mortality

Paralleling the negative effects on incarceration, keeping kids "out of trouble" during the summer could lead them down a safer path, and in extreme cases could even keep them alive. We observe in the IRS data that 0.38 percent of the sample of SYEP applicants die by

October 2014. In Table 7, we create a dummy representing whether an individual has died by 2014 in the IRS data and again estimate a linear probability two-stage least squares model.[30]

In the full population, SYEP reduces the probability of mortality by 0.08 percentage points (p=0.016). This represents a reduction in mortality of 19.92 percent relative to the baseline rate. In combination with the number of SYEP participants, the estimates imply a reduction of 86 deaths by 2014 due to the SYEP program in years 2005 to 2008.

While the small number of deaths prevent us from finding statistically significant differences in the treatment effect across groups, the absolute value of the point estimate is larger among males than among females; among Latinos, blacks, and other races than among whites; among the younger group than among the older group; among those who did not work prior to SYEP participation than among those who did work (paralleling the larger earnings crowdout among those who previously worked and those in the older group); and among those living in EZ's relative to other areas. These effects need not operate through a reduction in incarceration, and those whose lives were saved could differ from those kept from incarceration; nevertheless, the subgroups that show larger mortality effects typically, but not always, correspond to the subgroups that show an incarceration effect.

We also show the effect of SYEP on a dummy for whether an applicant died by Year 4 or Year 9 (Appendix Tables 17 and 18 show the effects separately by each year).[31] The cumulative effect is insignificant by Year 4 but becomes substantial and significant by Year 9. In data on the full U.S. population from the Social Security Administration actuarial life tables, as well as in our data, the yearly death rate increases several-fold from the mid-teen years to the mid-twenties. Thus, it is not surprising that we estimate larger effects on mortality in later years. These later mortality benefits indicate lasting effects of the program. The effects are insignificantly different across providers (p=0.22).

Cause of death

[30] Appendix Table 19 additionally shows that we obtain similar results with controls, with a Cox proportional hazard model, and with a probit.

[31] As death records enter the IRS data more promptly than earnings data, we have complete mortality data for the 2005 cohort through October 2014, which is in Year 9, whereas we have complete earnings and other data for the 2005 cohort only through Year 7.

As a secondary analysis, we investigate the particular causes of death that were affected by SYEP participation using DOH data. The DOH data only contain deaths in NYC and only cover years through 2013, whereas many of the deaths in the IRS data (among both SYEP lottery winners and losers) occur in 2014. Thus, we might expect to find smaller effect sizes, estimated with less statistical power, in the DOH analysis.

Given that our evidence suggests that SYEP "keeps kids out of trouble", we may be particularly interested in the effect of SYEP on the probability of death by "external causes," which include homicide, suicide, accidents, and other extrinsic causes.[32] These account for 69.61 percent of all deaths in our data by 2013 Among SYEP applicants in our sample, the most common cause of death is homicide, accounting for 47.52 percent of all deaths, and reflecting a much higher percentage than in the population as a whole in this age range.[33]

Table 7 shows our results. The DOH data reveal that SYEP causes a reduction in deaths from external causes (p=0.048), representing a 17 percent reduction relative to the control group.[34] The point estimate is 79 percent as large as the point estimate of the effect on mortality by any cause in the DOH data. The point estimate shows that SYEP reduces the probability of death by homicide, representing a 22 percent reduction relative to the control group, but the estimate is insignificant at conventional levels (p=0.15). The point estimate of effect on the probability of death by homicide is 64 percent as large as the effect on mortality by any cause in Row A. The effect on death from homicide is significant at 5 percent in various subgroups, including those who had not previously worked. When we estimate the regression for non-homicide external causes or "natural" causes, respectively, the point estimates are much smaller and are insignificant (p=0.25 and p=0.44, respectively).

To interpret the effect on dying in NYC, and of independent interest, we investigate whether SYEP has an effect on the probability that participants remain in NYC. Among those who work in Year 4, 88.87 percent work in NYC. Regressing a dummy for working in a NYC zip code in Year 4 on the SYEP participation dummy (instrumented as usual) shows a coefficient of -0.0028 on the SYEP dummy (p=0.11, insignificant at conventional levels).

---

[32] See http://www.nyc.gov/html/doh/downloads/pdf/vs/vs-population-and-mortality-report.pdf for the classification of causes into these categories. Natural causes represent all deaths other than external causes.
[33] See http://www.cdc.gov/injury/wisqars/pdf/10LCID_All_Deaths_By_Age_Group_2010-a.pdf
[34] Again, if treatment observations had better data quality (*i.e.* were more likely to match) than control observations, this would bias us toward estimating that SYEP *increases* mortality—the opposite of our findings.

SYEP likewise has no significant effect on the probability of working in NYS (relevant to the incarceration results above).

10. Conclusion

We investigate the effects of summer employment on youth by analyzing the New York City Summer Youth Employment Program from 2005 to 2008, which used a random lottery to select applicants for access to the program. We can now revisit the three broad rationales for programs that support summer youth employment: (1) transferring to youth; (2) raising future earnings, employment, or education; and (3) keeping youth "out of trouble."

We find support for the first rationale. SYEP increases contemporaneous employment and net earnings and transfers net income to participants. SYEP shows modest (19.24 percent) crowdout of other contemporaneous earnings, which is small relative to most previous results (Card, Kluve, and Weber 2010). Crowdout of earnings in the year of SYEP participation along with the subsequent years is more substantial (54.02 percent).

We find no evidence in favor of the second rationale. On balance, we find the opposite: SYEP lowers subsequent earnings for three years following SYEP participation, has little impact on the probability of future employment, and has no impact on college enrollment. The impact on subsequent earnings is small relative to likely lifetime income, but it is substantial relative to the size of the SYEP transfer (36.13 percent of the transfer).

Finally, we find that SYEP succeeds in the goal of keeping youth "out of trouble." SYEP leads to decreases in incarceration and mortality rates that are small in percentage point terms but large relative to the baseline rates. The reductions in incarceration and mortality parallel the typically positive effects on lower quantiles of the earnings distribution, suggesting that SYEP improves the left tail of outcomes, while the effects on higher quantiles of earnings suggest modest negative effects on the right tail. Similarly, the mortality and incarceration reductions typically appear strongest in the more disadvantaged groups, whereas the earnings crowdout typically appears largest for less disadvantaged groups. The stronger incarceration results in the 19 and older group for which we observe results for contemporaneous criminal activity suggest that a substantial portion of the effect on crime could operate through incapacitation (preventing youth from engaging in crime during the summer they are working) or other near-term effects on criminal activity. By contrast, the

mortality effects are significant only several years after SYEP participation — suggesting that in this case youth are kept "out of trouble" by putting them on a path that affects them years after the program.

The mortality effects have large benefits: the value of a statistical life (VSL) is estimated to be in the range of $9 million for prime-age workers (Viscusi and Aldy 2003, in $2013), implying benefits of $773.38 million for the estimated 86 lives saved by SYEP.[35] It is clear that the mortality benefits will be large within any plausible range of the value of life, although there is some uncertainty about the exact value of the benefits.[36] Meanwhile, the reduction in incarceration has more modest aggregate benefits: combining Donohue's (2009) estimates of the per-crime cost of an Index I crime with our estimates of the reduction in incarceration, the reduction in incarceration corresponds to a $4.66 million net benefit to society using Donohue's upper-end estimates of the benefits per crime, and a $1.03 million net benefit using Donohue's lower-end estimates.[37] Nonetheless, it is possible that the effect on crime could have much larger implications for the cost-benefit analysis if we incorporated the effect on prevalent crime outcomes that are unobserved in our data. It is illustrative to note that there were 59 times more total reported violent and property crimes in NYS in 2012 than the number of newly incarcerated individuals in a NYS prison (FBI 2012; Carson and Golinelli 2013). This suggests that the benefits of the impact on crime could be many times as large as the benefits above of reduced incarceration, if the impact on other crime outcomes is comparable to the impact on incarceration.[38] In future work, we hope to overcome the obstacles that we have so far encountered in accessing data on arrests of SYEP applicants.

---

[35] The lower end of the confidence interval shows that SYEP saved only eight lives, but all of the effects (including the earnings effects) are estimated with error — and the point stands that SYEP has mortality benefits that are likely quite large and therefore have the potential to be pivotal in the cost-benefit analysis.

[36] The lower end of the plausible range of the VSL is around $5.25 million (Viscusi and Aldy 2003), which would still imply very large SYEP mortality benefits of $451 million. For SYEP participants, who have more years of life remaining than the typical prime-age worker does, the VSL could be higher than $9 million (Viscusi and Aldy 2003). On the other hand, the value-of-life estimates typically have positive income elasticities, whereas SYEP participants typically have low income.

[37] Index I crimes include willful homicide, forcible rape, robbery, burglary, aggravated assault, larceny over $50, motor vehicle theft, and arson.

[38] However, Heller (2014) estimates a significant impact on violent crime arrests (for which eventual incarceration is more likely) but no impact on other arrests, raising the possibility that the impacts on other crime outcomes are not commensurate with the impact on incarceration. On the other hand, among the 19-and-older group, for which we have the best data, the incarceration benefits are much larger relative to the costs.

There are many costs and benefits we do not observe, and it is not possible to determine with certainty whether the benefits of the program outweigh the costs. For example, we do not observe the value of the goods produced by SYEP participants, the cost of other public programs, and so on. We are also unable to take into account any general equilibrium effects of SYEP.[39] It is possible that SYEP jobs displace jobs that employers would have otherwise offered, creating additional costs of the program — though we consider it unlikely that there is one-for-one displacement given that NYC pays for SYEP jobs.[40] We also do not observe externalities or non-pecuniary benefits of jobs. While we do not observe underground earnings, we believe it is unlikely that this would dramatically impact our results. To overturn our finding that SYEP decreased total earnings in Years 1-4, we would have to posit that SYEP raised underground earnings in these years. If anything, one might instead expect SYEP to push individuals into the formal sector, as SYEP itself is in the formal sector and our other evidence shows that SYEP industries are "sticky."

Nonetheless, it is clear that the $773.38 million in mortality benefits is substantial compared to plausible estimates of the costs of the program. Due to the SYEP program in the years 2005 to 2008, the discounted value of the reduction in non-SYEP earnings is $99.8 million; the discounted administrative costs of SYEP are $50.4 million, which is equal to the opportunity cost of these expenses if they are bought at competitive prices; theory tells us that the opportunity cost of time of SYEP participants should have been less than the discounted transfers to SYEP participants, or $186.0 million; and the deadweight cost of the taxes raised to fund SYEP equals the discounted accounting cost of SYEP, $236.4 million, multiplied by the marginal social cost of public funds.[41] While we cannot say with certainty whether SYEP's benefits outweigh its costs, it is clear that SYEP's mortality benefits are very large, and that they have a strong potential to be pivotal in determining whether the program's benefits outweigh its costs. Our results also suggest that earnings crowdout may

[39] Crepon et al. (2013) find that positive effects of job assistance come at the expense of other labor market participants. While it is possible that such general equilibrium effects could arise in our context, note that SYEP reduced individuals' subsequent earnings (which is unlikely to have come at the expense of others), and that SYEP is small relative to the entire NYC labor market and even the NYC youth labor market.

[40] In another context, a public employment program could increase the equilibrium wage and therefore cause displacement, but in our context wages are regulated to be at least the minimum wage. The crime or mortality impacts could also have general equilibrium effects, such as displacement by other crimes (*e.g.* Yang 2008).

[41] We discount to 2005 using a 3 percent real discount rate and express all dollar values in real 2013 terms.

be minimized, and the mortality reductions maximized, by targeting SYEP toward groups with weaker alternative job opportunities, such as younger individuals.[42]

As in other empirical settings, our estimates are local — in our case, to the SYEP compliers. However, our results may have important implications for other efforts to improve youth employment outcomes, including the Work Opportunity Tax Credit.[43] Indeed, the most salient difference between our study and previous work appears to be driven by our ability to observe mortality, as opposed to differences in sample characteristics. Like previous studies (*e.g.* Hollister, Kemper, and Maynard 1984, Couch 1992, Cave *et al.* 1993, Bloom *et al.* 1997), we find that SYEP did not increase future earnings and that earnings effects on their own could not justify the program's costs in a cost-benefit analysis. SYEP applicants are on average younger than in the programs studied in this previous work, and older SYEP applicants show the biggest negative earnings impacts. This suggests that we could observe even more negative earnings effects if our sample had the same age distribution as programs studied in this previous work.[44] Unlike previous studies, we find that the mortality reduction implies a very large source of new benefits, which has a strong potential to be pivotal in the cost-benefit analysis. SYEP participants tend to be disadvantaged, but unlike the programs analyzed in this previous work, which targeted youth with a record of delinquency or crime and/or youth who had dropped out of high school, SYEP does not include or exclude youth based on criminal records. The fact that the net benefits of SYEP tend to be largest in disadvantaged groups could suggest that the net benefits could be more positive if SYEP were targeted at such groups, thus reinforcing our conclusions.

Other active labor market programs for youth may or may not have such mortality benefits, but it is worth noting that like SYEP, other youth programs have been found to keep youth out of trouble (Schochet, Burghardt, and McConnell 2008; Heller 2014). Data should be gathered to determine whether mortality benefits appear in other contexts.

---

[42] In the younger group, the mortality gains are valued at $564 million; incarceration gains between $1 million and $3 million; the discounted earnings reduction at $29 million; discounted administrative costs at $25 million; the maximum opportunity cost of time is $75 million; and the deadweight cost is $100 million multiplied by the marginal social cost of public funds. The incarceration benefits may be largest in the oldest group, though that is difficult to judge given the data limitations. The results also suggest larger net benefits in EZs than other areas.

[43] Our results are consistent with previous studies finding limited earnings effects of subsidies for hiring disadvantaged groups like the WOTC (Hamersma 2003). If the WOTC leads to no significant positive impact on earnings, then the direct fiscal impact is not offset by changes in later tax receipts.

[44] The racial composition of youth are similar between these other programs and SYEP.

References

Angrist, Joshua D. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors." *Journal of Business & Economic Statistics* 19.1 (2001): 2-28.

Autor, David H., and Susan N. Houseman. "Do Temporary-Help Jobs Improve Labor Market Outcomes for Low-Skilled Workers? Evidence from 'Work First'." *American Economic Journal: Applied Economics* 2.3 (2010): 96-128.

Bellotti, Jeanne, Linda Rosenberg, Samina Sattar, Andrea Mraz Esposito, and Jessica Ziegler. *Reinvesting in America's Youth: Lessons from the 2009 Recovery Act Summer Youth Employment Initiative.* Princeton, NJ: Mathematica Policy Research, Inc, 2010.

Bloom, Dan, and Charles Michalopoulos. *How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research.* New York: Manpower Demonstration Research Corporation, 2001.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, and Johannes M. Bos. "The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study." *Journal of Human Resources* 32.3 (1997): 549-76.

Bushway, Shawn, and Robert Apel. "A signaling perspective on employment-based reentry programming." *Criminology and Public Policy* 11.1 (2012): 21–50.

Card, David, and Dean R. Hyslop. "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers." *Econometrica* 73.6 (2005): 1723-70.

Card, David, Jochen Kluve, and Andrea Weber. "Active Labour Market Policy Evaluations: A Meta-Analysis." *The Economic Journal* 120.548 (2010): F452-77.

Carson, Ann E., and Daniella Golinelli. *Prisoners in 2012: Trends in Admissions and Releases, 1991-2012.* Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics, December 2013.

Cave, George. *JOBSTART. Final Report on a Program for School Dropouts.* New York: Manpower Research Demonstration Corporation, 1993.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics* 125.1 (2010): 215-61.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. "How does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126.4 (2011): 1593-1660.

Cook, Philip J., Songman Kang, Anthony A. Braga, Jens Ludwig, and Mallory O'Brien. "An Experimental Evaluation of a Comprehensive Employment-Oriented Prisoner Re-entry Program." *Journal of Quantitative Criminology* (2014): doi 10.1007/s10940-014-9242-5.

Couch, Kenneth A. "New Evidence on the Long-Term Effects of Employment Training Programs." *Journal of Labor Economics* 10.4 (1992): 380-8.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment." *The Quarterly Journal of Economics* 128.2 (2013): 531-80.

Donohue III, John J. "Assessing the relative benefits of incarceration: Overall changes and the benefits on the margin." *Do prisons make us safer? The benefits and costs of the prison boom* (2009): 269-342. Ed. Steven Raphael and Michael Stoll. New York: Russell Sage Foundation.

Farkas, George, Randall Olsen, Ernst Stromsdorfer, Linda Sharpe, Felicity Skidmore, D. Alton Smith, and Sally Merrill. *Post-Program Impacts of the Youth Incentive Entitlement Pilot Projects*. New York: Manpower Demonstration Research Corporation, 1984.

Federal Bureau of Investigation. *Crime in the United States, 2012.* Washington, DC: US Department of Justice, October 2013. Retrieved February 22, 2015, from <http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/tables/5tabledatadecpdf/table_5_crime_in_the_united_states_by_state_2012.xls >.

Gelber, Alexander, Adam Isen, and Judd Kessler. "The Effects of Youth Employment: Evidence from New York City Summer Youth Employment Program Lotteries." NBER Working Paper 20810.

Grossman, Jean Baldwin, and Cynthia Sipe. *The Long-Term Impacts of the Summer Training and Education Program*. Philadelphia, PA: Public/Private Ventures, 1992.

Hamersma, Sarah. "An Evaluation of the Work Opportunity Tax Credit: Participation Rates and Employment Effects." *National Tax Journal* 56.4 (2003): 725-38.

Heckman, James J., Robert J. Lalonde, and James A. Smith. "The Economics and Econometrics of Active Labour Market Programs." *Handbook of Labour Economics, Volume 3A.* Eds. O. Ashenfelter and D. Card. Amsterdam and New York: Elsevier, 1999. 1865–2095.

Heller, Sara B. "Summer jobs reduce violence among disadvantaged youth." *Science*, Vol 346, 5 December 2014.

Hendra, Richard, James A. Riccio, Richard Dorsett, David H. Greenberg, Genevieve Knight, Joan Phillips, Phillip K. Robins, Sandra Vegeris, Johanna Walter, and Aaron Hill. *Breaking the Low-Pay, No-Pay Cycle: Final Evidence from the UK Employment Retention and Advancement (ERA) Demonstration*. New York: Manpower Demonstration Research Corporation, 2011.

Hollister, Robinson G., Peter Kemper, and Rebecca A. Maynard. *The National Supported Work Demonstration*. Madison, WI: Univ of Wisconsin Press, 1984.

LaLonde, R. "Employment and Training Programs." *Means-Tested Transfer Programs in the United States.* Ed. Robert Moffitt.University of Chicago Press, 2003.

Lee, David S. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76.3 (2009): 1071-102.

Leos-Urbel, Jacob. "What is Summer Job Worth? the Impact of Summer Youth Employment on Academic Outcomes." New York University, 2012.

McClanahan, Wendy, Cynthia L. Sipe, and Thomas J. Smith. "An Evaluation of the Summer Career Exploration Program." Public/Private Ventures, 2004.

Pallais, Amanda. "Inefficient Hiring in Entry-Level Labor Markets." Forthcoming, *American Economic Review* (2014).

Sabol, William J., Heather C. West, and Matthew Cooper. *Prisoners in 2008*. Washington, DC: U.S. Department of Justice, 2008.

Schochet, Peter Z., John Burghardt, and Sheena McConnell. "Does Job Corps Work? Impact Findings from the National Job Corps Study." *The American Economic Review* 98.5 (2008): 1864-86.

Stanley, Marcus, Lawrence F. Katz, and Alan B. Krueger. "Developing Skills: What We Know about the Impact of American Employment and Training Programs on Employment, Earnings and Educational Outcomes." (1998).

Statistical Abstract of the United States: 2000 (116th Edition). Washington, DC: U.S. Census Bureau, 1999.

Viscusi, W. Kip, and Joseph E. Aldy. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty* 27.1 (2003): 5-76.

Yang, Dean. "Can Enforcement Backfire? Crime Displacement in the Context of Customs Reform in the Philippines." *Review of Economics and Statistics* 90.1 (2008): 1-14.

Table 1. *Treatment-control balance.*

(1) Variable	(2) Mean (SD)	(3) Coeff. (SE) on Treatment
Main Outcomes (Years 0 to 4)		
Total yearly earnings	3,555.29 (7,201.94)	--
NYC gov't yearly earnings	218.39 (474.92)	--
Non-NYC gov't yearly earnings	3,336.91 (7,259.75)	--
Has any job	0.63 (0.48)	--
Has any non-NYC gov't job	0.50 (0.50)	--
College enrollment	0.23 (0.42)	--
Lagged Outcomes (Year -1)		
Total earnings	890.33 (4,489.23)	-23.65 (21.55)
NYC gov't earnings	256.96 (495.85)	-1.79 (2.13)
Non-NYC gov't earnings	633.37 (4,477.73)	-21.87 (21.32)
Has any job	0.32 (0.47)	-0.0024 (0.0018)
Has any non-NYC gov't job	0.13 (0.34)	-0.0017 (0.0012)
College enrollment	0.04 (0.20)	0.00086 (0.00069)
Family income	39,521.56 (29,412.29)	-36.99 (115.75)
SYEP participation	0.21 (0.41)	-0.0014 (0.0018)
Race		
White	0.13 (0.33)	-0.0019 (0.0014)
Latino	0.27 (0.44)	0.00067 (0.0015)
Black	0.48 (0.50)	0.00073 (0.0017)
Other	0.12 (0.33)	0.00053 (0.0016)
Other variables		
Male	0.45 (0.50)	-0.0025 (0.0022)
Age	16.50 (1.63)	0.00057 (0.0086)
# Family Members	4.30 (1.86)	-0.0023 (0.0068)
U.S. Citizen	0.93 (0.25)	-0.00068 (0.00098)
SYEP-IRS Match dummy	0.998 (0.06)	-0.00032 (0.00024)

Notes: The table shows summary statistics and demonstrates that there are no significant differences in covariates across the treatment and control groups. In Column 2, we report means of variables, with standard deviations in parentheses. In Column 3, we use OLS to regress the variable in question on a dummy for winning the SYEP lottery and provider-year fixed effects, and report coefficients and standard errors on the SYEP win dummy from this regression. 294,580 observations are included in the sample for all variables, except in the case of measuring prior year SYEP participation (238,023 observations). Main outcomes are observed in years 0-4 (inclusive) and are observed at a yearly level (so that, for example, the mean of the "has any job" dummy refers to the probability that an individual has a job in a given year). Lagged outcomes are observed in the calendar year prior to the SYEP lottery in question. Family income refers to income from SYEP lottery participants' tax unit. All outcomes are derived from IRS data except gender, race, citizenship, age, and SYEP participation, which are derived from SYEP administrative data. "--" indicates that for the main outcomes, readers should refer to subsequent tables, which investigate the effect of SYEP on these outcomes in detail. For binary outcomes, we report the mean of a dummy that equals 1 if the characteristic is observed. "Match dummy" refers to a dummy variable that equals 1 if the individual was matched to tax records according to SSN or gender, DOB, name, and first or last four digits of the SSN. *** denotes significance at the 10% level; ** at the 5% level; and * at the 1% level.

Table 2: *Effect of SYEP lottery win in Year 0 on SYEP participation in each year*

	(1) SYEP Participation	(2) F-statistic
A) Year 0	0.73	4188.68
	(0.011)***	
B) Year 1	0.031	81.90
	(0.0034)***	
C) Year 2	0.011	48.16
	(0.0016)***	
D) Year 3	0.0031	10.89
	(0.00093)***	
E) Year 4	0.0015	6.45
	(0.00059)***	

Notes: The table shows the results of OLS regressions in which SYEP participation in Year 0 and subsequent years is related to SYEP participation in Year 0, controlling for provider-by-year fixed effects. The table shows marginal effects and standard errors on the SYEP participation dummy. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 3. *Effect of SYEP participation on earnings and employment*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Any job	(5) Any non-NYC gov't job
A) Year 0	876.26	1085.34	-208.87	0.71	-0.048
	(25.08)***	(10.14)***	(24.83)***	(0.0063)***	(0.0035)***
	[1,153.07]	[1,112.36]	[40.71]	[0.30]	[0.27]
B) Year 1	-99.61	45.98	-145.59	0.012	-0.018
	(40.06)**	(5.01)***	(40.08)***	(0.0034)***	(0.0026)***
	[2,241.12]	[2,036.89]	[204.25]	[0.53]	[0.40]
C) Year 2	-94.16	23.29	-117.45	0.0045	-0.0097
	(42.05)**	(3.33)***	(42.42)***	(0.0031)	(0.0027)***
	[3,246.27]	[3,105.56]	[140.71]	[0.60]	[0.52]
D) Year 3	-110.81	8.22	-119.04	-0.00063	-0.0051
	(44.43)**	(2.08)***	(44.26)***	(0.0023)	(0.0023)**
	[4,471.66]	[4,380.98]	[90.69]	[0.66]	[0.62]
E) Year 4	-35.15	4.47	-39.62	0.0013	-0.00032
	(44.82)	(1.58)***	(45.23)	(0.0022)	(0.0022)
	[5,941.65]	[5,887.48]	[54.16]	[0.72]	[0.69]
F) Years 0-4	536.53	1167.30	-630.57	0.089	-0.0059
	(173.10)***	(15.09)***	(173.75)***	(0.0033)***	(0.0022)***
	[17,053.76]	[16,523.26]	[530.52]	[0.89]	[0.83]
G) Years 1-4	-339.73	81.96	-421.70	0.010	-0.0027
	(154.51)**	(9.51)***	(154.94)***	(0.0021)***	(0.0021)
	[15,900.70]	[15,410.91]	[489.81]	[0.88]	[0.82]

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions (1)-(2) of earnings and employment outcomes on SYEP participation. Standard errors are in parentheses, and control group means are in square brackets. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. Each row shows the results for a different year or set of years, and each column shows the results for a different outcome. Our measure of annual earnings comes from W-2s, mandatory information returns filed with the IRS by employers for each employee. Having "any job" is defined as earning a positive amount of income, as reported on the W-2 form. We control for SYEP provider-by-year dummies so that the estimates are driven by random variation in winning the SYEP lottery. In Year 0, the control group mean and the effect on the jobs dummy do not add to one; this is a consequence of the fact that there are separate lotteries in each provider in each year, with different numbers of observations in each lottery. The number of observations in each regression is 294,580, corresponding to 198,745 individuals. Standard errors are clustered at the level of the SYEP provider. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 4: *Effect of SYEP participation on earnings and employment by industry and job type*

	(1) Cluster 1	(2) Cluster 2	(3) For-profits	(4) Non-profits	(5) Gov't
Panel A: Effects on total earnings					
A) Year	966.59	-91.70	-75.56	786.65	165.13
0	(12.62)***	(24.96)***	(26.88)***	(28.29)***	(25.87)**
	[209.54]	[945.27]	[1,052.03]	[40.38]	[61.14]
B Years	983.43	-443.54	-435.77	854.81	117.06
0-4	(55.80)***	(164.61)***	(170.53)***	(38.33)***	(51.72)**
	[2,834.19]	[14,224.69]	[15,714.70]	[848.21]	[492.96]
C) Years	16.84	-351.85	-360.21	68.16	-48.07
1-4	(52.89)	(146.90)**	(152.81)**	(16.88)***	(34.26)
	[2,624.65]	[13,279.42]	[14,662.67]	[452.57]	[787.07]
Panel B: Effects on having any job					
A) Year	0.81	0.046	0.043	0.67	0.16
0	(0.0063)***	(0.0067)***	(0.0090)***	(0.022)***	(0.021)***
	[0.093]	[0.24]	[0.26]	[0.033]	[0.034]
B) Years	0.45	0.0085	0.011	0.51	0.13
0-4	(0.0067)***	(0.0025)***	(0.0025)***	(0.016)***	(0.017)***
	[0.46]	[0.81]	[0.83]	[0.23]	[0.20]
C) Years	0.034	0.00063	-0.00021	0.040	0.0049
1-4	(0.0039)***	(0.0022)	(0.0022)	(0.0039)***	(0.0029)*
	[0.44]	[0.79]	[0.82]	[0.21]	[0.18]

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from IV regressions of earnings and employment outcomes on SYEP participation. The mean of the dependent variable in the control group is shown in brackets, below the standard error in parentheses. The instrument for whether an individual participated in SYEP is a dummy indicating that the individual won the SYEP lottery. Columns 1-2 show the results of IV regressions in which earnings (Panel A) or the probability of having a job (Panel B) in a given industry cluster and year are related to SYEP participation. Using DYCD's industry classification, Cluster 1 corresponds to industries that are overrepresented among SYEP lottery winners relative to SYEP lottery losers: arts and recreation, camp (out of city), community/social service, day care/day camp, educational services, and healthcare/medical. We classify these as belonging to one of the following cluster of NAICS codes: 61, 62, 71, and 92. Cluster 2 corresponds to other SYEP classifications and NAICS codes. Columns 3-5 show the results of IV regressions in which earnings or the probability of having a job in a given sector (for-profit, non-profit, or government) are related to SYEP participation. In Years 0-4, mean yearly earnings in Cluster 1 is $681.08; in Cluster 2 is $2,875.31; in for-profits is $3,190.88; in non-profits is $187.45; and in government employers is $177.36. In Years 0-4 (considering each year as a separate observation), the probability of employment in Cluster 1 and Cluster 2 is 25.47 percent and 46.56 percent, respectively. In Years 0-4 (considering each year as a separate observation), the probability of employment in the for-profit, non-profit, and government sectors is 48.94, 7.63, and 13.56 percent, respectively. See Gelber, Isen, and Kessler (2015) for estimates in each year from Year 1 to Year 4. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level. See other notes to Table 3.

Table 5. *Effect of SYEP participation on college enrollment*

	(1) Coefficient (SE) on SYEP participation	(2) Dependent variable mean
A) Year 0	0.0011	0.076
	(0.0015)	
B) Year 1	0.0029	0.15
	(0.0019)	
C) Year 2	-0.0012	0.24
	(0.0024)	
D) Year 3	0.00049	0.33
	(0.0021)	
E) Year 4	-0.0032	0.36
	(0.0024)	
F) Years 0-4	-0.0035	0.50
	(0.0025)	
G) Years 1-4	-0.0029	0.49
	(0.0025)	
H) Total years of college	-0.000033	1.15
	(0.0071)	

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions of a college attendance dummy or total years of college on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. The results are similar if we limit the sample to those 18 years of age and older (because younger individuals are unlikely to go to college). Rows F and G investigate the impact of SYEP enrollment on the probability that an individual attends college at some point during Years 0-4 and 1-4, respectively. Row H shows the effect of Year 0 SYEP participation on the total number of years enrolled in college over Years 0-4 cumulatively. The mean total number of years enrolled in college over Years 0-4 is 1.17. Column 2 shows the mean of the dependent variable in the control group that lost the lottery. See other notes to Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 6. *Effect of SYEP on incarceration*

	(1) 2SLS	(2) Incarceration mean (x 100)	(3) N	(4) *p*-value of test for equality
(A) Full population	-0.098 (0.046)**	0.95	294,580	--
(B) 19 and older	-0.48 (0.22)**	0.85	24,809	0.12
(C) 16.25 to 18	0.0017 (0.076)	1.00	122,601	
(D) Under 16.25	-0.12 (0.071)	0.93	147,320	
(D) Males	-0.22 (0.094)**	2.01	132,692	0.011
(E) Females	0.023 (0.020)	0.083	161,888	
(F) White	-0.15 (0.087)*	0.073	37,172	0.16
(G) Black	-0.16 (0.070)***	1.48	142,627	
(H) Latino	0.049 (0.081)	0.70	79,095	
(I) Other	-0.062 (0.098)	0.33	35,686	
(J) Prior work	-0.054 (0.091)	0.79	94,855	0.51
(K) No prior work	-0.12 (0.050)**	1.02	199,725	
(L) Emp. Zone	-0.057 (0.080)	1.08	149,379	0.38
(M) Not Emp. Zone	-0.14 (0.051)	0.82	145,200	

Notes: Column 1 shows coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions of a dummy for incarceration in NYS on SYEP participation. Column 2 shows the mean of the dependent variable in the control group. Each row shows the results for a different population. We multiply the incarceration dummy by 100 so that coefficients show percentage point changes (for the reader's ease). Note that the probit specification runs "reduced form" regressions that regress the dependent variable directly on the dummy for winning the SYEP lottery, not an instrumental variables regression. The probit coefficients represent marginal effects, calculated at the mean. 16.25 represents the approximate median age in the sample. "Emp. Zone" refers to an Empowerment Zone. The final column reports *p*-values from tests of equality across the coefficients estimated across sub-groups within a given category (e.g. across race groups, or across ages, or across genders). *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 7. *Effect of SYEP on mortality*

	(1) 2SLS	(2) Mortality mean (x 100)	(3) N	(4) *p*-value of test for equality
A) Full population	-0.075 (0.031)**	0.38	294,238	--
B) Males	-0.15 (0.06)**	0.61	132,692	0.054
C) Females	-0.016 (0.032)	0.19	161,888	
D) White	0.0030 (0.076)	0.16	37,172	0.40
E) Black	-0.061 (0.048)	0.52	142,627	
F) Latino	-0.14 (0.055)**	0.32	79,095	
G) Other races	-0.054 (0.074)	0.18	35,686	
H) Older	-0.029 (0.047)	0.40	147,260	0.17
I) Younger	-0.11 (0.040)***	0.33	147,320	
J) Work in Year -1	0.034 (0.060)	0.35	94,855	0.048
K) No work in Year -1	-0.12 (0.043)***	0.39	199,725	
L) Emp. Zone	-0.090 (0.044)**	0.37	149,201	0.67
M) Not Emp. Zone	-0.063 (0.044)	0.39	145,037	
N) By Year 4	-0.015 (0.022)	0.19	294,238	0.0070
O) By Year 9	-0.22 (0.085)***	0.55	56,557	

Notes: Column 1 shows a coefficients and standard errors on the SYEP participation dummy from two-stage least squares estimate using a linear probability model. Each row shows the results for a different population (rows B through K) or a different time period (rows L and M). We eliminate from the regressions those rare cases of individuals who died prior to participating in SYEP, which explains the difference in the sample size between the full sample here and elsewhere, and is uncorrelated with winning the lottery. Column 2 shows the mean of the mortality dummy, multiplied by 100, in each group or time period. So that readers can more easily interpret the results, we have multiplied the dependent variable by 100. In rows L and M, we show the effect of SYEP on a dummy for whether an applicant died *by* a given year. Column 1 shows the results of our two-stage least squares specification (1)-(2). See Tables 3 and 6 for other notes and information on samples. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Table 8. *Effect of SYEP on mortality by cause of death.*

	(1) 2SLS	(2) Percent of deaths by 2012
Panel A: death from any cause comparison between DOH and IRS data		
A) Any cause (DOH)	-0.052	100
	(0.024)**	
B) Any cause (IRS through 2013)	-0.061	100
	(0.028)**	
Panel B: specific causes of death from DOH data		
C) External causes (DOH)	-0.041	69.61
	(0.021)**	
D) Homicide (DOH)	-0.027	47.52
	(0.019)	
E) Non-homicide external causes (DOH)	-0.015	22.09
	(0.013)	
F) Natural causes (DOH)	-0.011	30.39
	(0.014)	

Notes: Each row shows the results of a different regression where a dummy for a different cause of death is the dependent variable in our 2SLS, linear probability model (1)-(2). So that readers can more easily interpret the results, we have multiplied the dependent variable by 100. The results are comparable with Cox or probit models; we show a two-stage least squares model here to show results that are comparable to the IV results elsewhere in the paper. It is not surprising to find a slight discrepancy between the DOH and IRS data results in Rows A and B, because the IRS data cover all deaths whereas the DOH data cover only deaths in NYC. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Online Appendix 1: Other specifications

Robustness checks

Various robustness checks to the basic results on earnings in Table 3 yield extremely similar results. As we show in Appendix Table 5, all of the results are very similar to those in Table 3 when we perform the same specifications but additionally control for: gender, dummies for race categories, citizenship, age, number of family members, individual's wage income in Year -1, individual's NYC government wages in Year -1, individual's non-NYC government wages in Year -1, a dummy for whether the individual had any job in Year -1, a dummy for whether the individual had a non-NYC government job in Year -1, a dummy for whether the individual was enrolled in college in Year -1, a dummy for whether the individual was claimed on a return in Year -1, and family income in Year -1 if the individual was claimed on a return in Year -1.[i]

Appendix Table 6 shows a number of other specifications, all of which deliver results very similar to the baseline specification. First, in Panel A, we use the dummy for winning the *initial* SYEP lottery as the instrument, rather than our baseline where the instrument is a dummy for winning any of the SYEP lotteries in Year 0.[ii] Second, in Panel B, we only include individuals in the sample who match the SYEP data according to their SSN. Third, in Panel C, we cluster at the individual level rather than at the level of the provider, which leads to extremely similar standard errors and significance levels.

Dynamics

As shown in Table 2, SYEP participation in Year 0 slightly affects the probability of SYEP participation in future years (i.e. Years 1-4). Thus, some of the effects on earnings we observe are mediated through the impact of SYEP on future SYEP participation, though the small effect on future participation suggests a limited role for such a mechanism. To more precisely examine the extent to which this drives the results, we estimate the effect of SYEP participation on earnings using the "dynamic" specification of Cellini, Ferreira, and Rothstein (2010). In our context, this dynamic estimator effectively yields the effect of SYEP participation in Year 0 on earnings in any subsequent year, while removing the effect that operates through the channel of the effect of Year 0 SYEP participation on subsequent SYEP participation. By contrast, the instrumental variables estimates in Table 3, which we call the "static" estimates, estimate the effect of being employed through SYEP in Year 0 on the future path of earnings and employment, including the effect that works through future SYEP participation.

Following the "recursive" procedure of Cellini et al. (2010), we first estimate the coefficients in Tables 2 and 3, showing the effect of Year 0 SYEP participation on subsequent earnings and on subsequent SYEP participation respectively, using the

[i] Controlling for higher-order terms in income also has negligible affects on the results.
[ii] Unsurprisingly, these regressions have somewhat larger standard errors, but the estimates are still very significant and have small confidence intervals.

methods discussed above. Let β_τ^S represent the estimate of the effect of SYEP participation in Year 0 on earnings (or another outcome variable) in year τ from Table 3, and let π_t represent the effect of Year 0 SYEP participation on the probability of SYEP participation in year *t* (from Table 2). We calculate the dynamic effect β_τ^D in year τ as:

$$\beta_\tau^D = \beta_\tau^S - \sum_{t=1}^{\tau} \pi_t \beta_{\tau-t}^D \qquad (3)$$

We solve for the dynamic effects in each year using the recursive equation (3). Standard errors are obtained by the delta method. By contrast, the instrumental variables estimates in Table 3, which we call the "static" estimates and which represent β_τ^S in (3), estimate the effect of being employed through SYEP in Year 0 on the future path of earnings and employment. These two objects of study reflect different conceptual experiments, both of which are of interest.

Since Table 2 shows that SYEP participation in Year 0 has a small impact on the probability of future SYEP participation, it is not surprising that the dynamic estimator finds results that are similar to the static estimates. Appendix Table 7 shows that the effect on subsequent total earnings is somewhat more negative in the dynamic specification than in the static specification, particularly in the initial years. Nonetheless, it is worth noting that the estimated effect on non-NYC government earnings and on total earnings is generally similar to the estimated effect in the static specification in Table 3.

The small positive effect of SYEP participation on NYC government earnings in Years 1-4 in the dynamic specification indicates that average total wages conditional on SYEP employment in subsequent years must be increasing slightly over time, likely because of the rise in the minimum wage over time (and possibly because average hours worked in SYEP could have changed).

In our other analysis throughout the paper, we use the simpler static specification, though we note that throughout all of our specifications and outcomes we obtain comparable results in the dynamic version to those in the static analysis (which is unsurprising since the effect of Year 0 SYEP participation on subsequent SYEP participation is quite modest).

Appendix Table 8 shows specifications relating to the number of times the individual participated, or could have participated, in SYEP. Among applicants who were too young to have been eligible to participate in SYEP previously because they were 14 or younger in 2005 (115,337 applicants), the results are similar to our main sample.[iii] While we only have records from SYEP on lottery applications and SYEP participation starting in 2005, we can use IRS records on NYC government earnings in prior years as a proxy for prior SYEP participation (classifying individuals as having participated in SYEP in a given prior year when they had positive NYC government earnings in that year). Under this

[iii] Among those 21 years old in the year of SYEP application, who are ineligible to participate in SYEP again, the sample size drops to 16,620, and the results are imprecise and insignificant (p>0.40).

definition, we show the effect of SYEP participation separately for those who had previously participated in SYEP no times, one time, two times, three times, or four or more times. This is relevant because our empirical specification assumes that the effect of moving from zero times participating in SYEP to one time participating in SYEP is the same as the effect of moving from one time participating to two times, and so on. The effects in these groups are statistically indistinguishable. We also use winning the SYEP lottery as an instrument for the total number of times participating in SYEP between 1999 and Year 0 (inclusive) and show that this also yields comparable results.

Online Appendix 2: Discussion of Mechanisms

It is worth considering why SYEP participation reduced mean earnings for three years after participation in the program. In this Appendix, we consider a number of additional potential explanations, including: replacing work experience; job type; effects on job transitions; income effects; substitutability of leisure across years; and signaling.

Replacing valuable work experience

As noted in the main text, it is possible that SYEP harms individuals' future earnings because it affects the experiences that they gain in Year 0. If SYEP has a negative effect on subsequent earnings in part because it crowds out other, valuable employment experiences in Year 0, then we would expect that groups with more Year 0 crowdout would also show more negative effects of SYEP on subsequent earnings. This is borne out in the data.

The groups for which the effect on total earnings in Year 0 is smallest are the groups for which the effect on non-NYC government earnings in Year 0 is particularly negative (i.e. groups for which the crowdout due to participating in SYEP is particularly large). In these groups, the effect on earnings in Years 1-4 also tends to be particularly negative, consistent with the hypothesis that such crowdout plays a role in explaining the negative effect of SYEP on subsequent earnings.[iv]

This pattern holds true for all the groups we analyze. First, the effect of SYEP participation on total earnings in Year 0 is highest in the 50th percentile, intermediate in the 75th percentile, and smallest in the 90th percentile. The effects on subsequent earnings are ordered the same way: the effect on earnings in Years 1-4 is sometimes positive in the 50th percentile, negative in the 75th percentile, and most negative in the 90th percentile. Similarly, the point estimate of the effect of SYEP participation on non-NYC government earnings in Year 0 among whites is particularly negative (-$295.99), is intermediate among Latinos ($-253.94), and is least negative among blacks and others (-$176.41 and -$194.20). This correlates with the effects on subsequent earnings estimated in each

[iv] Nonetheless, as noted in the main text, we emphasize that each of these samples tends to differ on average along many characteristics (that are correlated across samples), and it could be that the effect on non-NYC government earnings in Year 0 is correlated across groups with the effect on subsequent total earnings for reasons unrelated to the hypothesis described above.

group, which is most negative for whites, intermediate for Latinos, and least negative for blacks and others. Likewise, the effect on non-NYC government earnings in Year 0 is more negative among the older SYEP applicants (-$384.56) than among the younger applicants (-$67.08), and there is a larger negative effect on subsequent earnings in the older group. Moreover, the effect on non-NYC government earnings in Year 0 is much more negative for those who worked prior to SYEP (-$401.29) than that among those who did not work prior to SYEP (-$117.76), and those who worked prior to SYEP showed the more negative effect on subsequent earnings. Finally, the Year 0 point estimate on non-NYC government earnings is much more negative in the 2007-8 lotteries than in the 2005-6 lotteries, and the effect of SYEP on future earnings is much more negative in the 2007-8 lotteries. It is also noteworthy that more negative subsequent impacts on earnings tend to occur in groups that are more likely to otherwise have a job in Year 0, like the older group or those who had a job in the prior year.

Effects of type of job

We investigate whether the type of SYEP job individuals are placed in has implications for their earnings. In Appendix Table 12, we show the following OLS regressions:

$$E_{ij} = \beta_0+\beta_1(W_{ij}*P_j)+\beta_2 W_{ij}+X_j\beta+v_{ij} \quad (4)$$

where E_{ij} represents the earnings of individual i in lottery j, W_{ij} is a dummy for winning the SYEP lottery, P_j is the percent of a provider's jobs that are in Cluster 1 (in a given lottery j), and X_j reflect dummies for each provider-lottery combination.[v] Thus, the coefficient β_1 on the interaction term reflects whether individuals who win the lottery at providers with a greater proportion of Cluster 1 jobs have lower or higher earnings than those who win the lottery in providers with a smaller proportion of Cluster 1 jobs.[vi] (If an individual applies to a provider with a greater proportion of Cluster 1 jobs, then winning the lottery is more likely to place the individual in a Cluster 1 job.)

The results in Panel A of Appendix Table 12 show that being in a Cluster 1 job negatively impacts subsequent total earnings (significantly in Year 3 and Year 4, as shown in Column 1), and that this is driven by negative effects on earnings in Cluster 2 (as shown in Column 3). Thus, the regressions demonstrate that placing people in jobs in SYEP-type industries (i.e. Cluster 1 industries) has a substantial negative impact on earnings in other industries, while having no significant impact on earnings in SYEP-type industries. This is evidence that SYEP affects future earnings in part because it affects the type of job that individuals take in future years.

In Panel B of Appendix Table 12, we investigate the coefficient on the main effect of winning the SYEP lottery. This coefficient reflects the hypothetical impact of winning the SYEP lottery *in a provider that only places individuals into Industry Cluster 2.*

---

[v] The IV version of this regression — using W_{ij} and $W_{ij}*P_j$ as instruments for SYEP participation and SYEP participation in provider P_j — shows directionally similar results with less statistical power.

[vi] We also explored regressions in which we interacted winning the lottery with the percent of jobs at the provider that were private not-for-profit, private for-profit, or government jobs. We found no significant differences in the effects of SYEP across these groups.

Intriguingly, the effect on total earnings is positive and significant in many cases (specifically in Year 0, Year 3, Year 4, Years 0-4, and Years 1-4), and it is positive and insignificant in the remaining cases. The point estimates are substantial (several hundred dollars in the case of individual years, and several times larger in the case of Years 0-4 or Years 1-4 combined). If SYEP has a positive effect on earnings when individuals are in Industry Cluster 2, then SYEP could improve outcomes of its lottery winners by increasing the fraction of SYEP jobs that are in Industry Cluster 2.[vii]

Nonetheless, we reiterate the important caveat that heterogeneity in the effects across providers could be driven by factors that happen to be correlated with the types of jobs in each provider. In this case, these results could not be interpreted as the causal effects on earnings of Cluster 1 vs. Cluster 2 jobs.

The effect of SYEP on subsequent job industry does not account for the effect of SYEP on subsequent earnings: when we estimate the effect of SYEP on the probability of working in each two-digit industry and calculate the mechanical effect of this industry pattern on earnings (using mean earnings in each industry and year among the control group), this accounts for an insignificant fraction of SYEP's effect on future earnings.

Effects on employment transitions

It is also possible that SYEP harms individuals' career development by increasing employment transitions and interrupting experience with past employers.[viii] Youth could take the SYEP job rather than continuing with an existing employment relationship. For example, a youth who had worked at a summer job in the previous summer might choose not to return to the same employer if a SYEP job were available. Appendix Table 11 shows that SYEP has such an effect in Year 0 (among those who did not participate in SYEP in Year -1), though this effect is small. In Year 1, the estimate is also negative but barely significant, and in subsequent years the estimates turn insignificant.

Employment transitions are associated with lower earnings in a cross-section of individuals, and when we use the size of this cross-sectional association, we find that the effect of SYEP on job transitions can account for 38 percent of the impact of SYEP on earnings in Years 1 to 4. However, we emphasize that the cross-sectional association between job transitions and earnings is not causal and is therefore subject to omitted variable bias. We also note that like other channels explored in this appendix, this channel is not mutually exclusive with others we have explored.[ix]

[vii] We also note the caveat that when we control for all other available demographics interacted individually with W_{ij} (which adds many controls and should therefore reduce the efficiency of the estimates), the coefficients β_1 and β_2 above are reduced in significance and substantially reduced in magnitude, although we robustly estimate that β_1 is negative and substantial and that β_2 is positive and substantial.

[viii] Card and Hyslop (2005) examine a related issue when investigating the dynamic effects of the Self Sufficiency Project in Canada.

[ix] If this is the primary explanation for negative effects on future earnings, we also would not immediately expect the variation across provider industries that we find.

Income effects or substitutability of leisure

One potential explanation for the decrease in subsequent earnings relates to income effects. Getting a SYEP job leads to an average increase in earnings of $872.36 in the year of the SYEP job, which could in principle lead to increased leisure in subsequent years if leisure is a normal good. However, income effects cannot immediately explain the striking heterogeneity across groups that we find; for example, we would have to postulate that there is an income effect on the earnings of those who previously had a job, but not on the earnings of those who previously did not have a job. While it is possible that the income effects are heterogeneous in ways that track the heterogeneous findings across groups, this is an *ad hoc* — and not particularly parsimonious — explanation. Moreover, recall that SYEP leads individuals to earn less conditional on having a job. Consequently, such an income effect would have to operate in a manner that seems unexpected: it would have to decrease earnings even as it leads individuals to be equally or more likely to take a job.

Moreover, as we discuss in further detail below, one of our robust findings is that individuals in groups that experienced larger increases in total earnings in Year 0 also experienced smaller earnings decreases in subsequent years. If an income effect were responsible for the results and were homogeneous across groups, we might have expected the opposite (assuming that leisure is a normal good). Note, however, that it is possible that in Year 0, the increase in income due to SYEP caused a decrease in non-SYEP earnings in Year 0 *subsequent* to SYEP participation (i.e. in the fall of Year 0).

In principle, another explanation for the results is that leisure could be substitutable across years, so that a decrease in leisure in Year 0 would have been associated with an increase in leisure in Year 1. However, this explanation runs into the same set of difficulties as the income effect explanation just explored. Again, leisure substitutability cannot immediately explain the striking heterogeneity across groups that we find; such an explanation would be *ad hoc*. Again, such leisure substitutability would not be consistent with the finding that the groups with larger increases in Year 0 total earnings tended to be those with smaller subsequent decreases in total earnings. And again, such leisure substitutability would be operating in a way that seems unexpected: it would have to decrease earnings even as it leads individuals to be more likely to take on a job. Again, a more satisfying hypothesis could explain both the effect of SYEP on the probability of having a job and the effect of SYEP on subsequent earnings.

Signaling

Another possible explanation is that employers use the information that an individual participated in SYEP in deciding whether to hire them. While winning the SYEP lottery is random conditional on applying, SYEP participation still contains information that employers could use. Those who apply for and enroll in SYEP may be those who have difficulty securing employment elsewhere. Thus, SYEP participants may be negatively

selected relative to the population as a whole.[x] Employers may therefore take the fact that an individual participated in SYEP as a negative signal of their productivity (in contrast to receiving a positive signal from the employment of an otherwise inexperienced worker, as in Pallais 2014).[xi] This would be consistent with some of the patterns across groups: groups with lower income on average (like blacks or younger individuals) tend to have negative effects on subsequent earnings that are smaller in absolute value, which we might expect if SYEP participation is interpreted less negatively in more disadvantaged groups (because more disadvantaged groups are less likely to have alternative options).

However, the signaling explanation is not immediately consistent with the difference in effects before and during the Great Recession. The point estimates of the effects on Years 1-4 earnings in the lotteries during the Great Recession (2007 and 2008) are more negative than those in the lotteries prior to the Great Recession (2005 and 2006). If employers were updating their expectation of individuals' productivity on the basis of SYEP participation, one might expect that employers would interpret SYEP participation more negatively when the individual participated before the Great Recession, than when the individual participated during the Great Recession (since it was more difficult to find other employment during the Great Recession). At the same time, the estimates in 2005-2006 are insignificantly different from those in 2007-8, though barely so. Thus, while our evidence is not directly inconsistent with the signaling hypothesis, it also does not fully support the signaling hypothesis either.

Other potential explanations

It is also worth mentioning a number of other possible explanations for the results. First, it is possible that SYEP caused the labor supply curve to shift to the right: SYEP could lead individuals to be more willing to accept low-paying jobs of the sort SYEP offers. If the demand curve shifted to the left, we would expect a decrease in earnings and a decrease in hours worked. If the supply curve shifted to the right, we would expect an increase in hours worked and could observe a decrease or increase in earnings. Table 3 shows that total jobs increase in Year 1, although some of this increase is due to an increase in SYEP jobs, which do not reflect labor demand. The dynamic estimates effectively remove the influence of SYEP jobs and show that total jobs decrease in Year 1. While we lack a measure of hours worked for the full sample, employment may be a reasonable proxy for hours worked given the absence of other proxies.

Second, SYEP participants could be exposed to a peer group that has negative effects on their future earnings. We find no evidence for such a channel: when we interact winning

[x] Indeed, our data show that (unconditional on SYEP application) prior family income of SYEP participants is substantially lower than that of those who were eligible on the basis of being NYC residents but did not participate in SYEP.

[xi] SYEP has negative earnings effect for those who had previous employment, but no significant negative earnings effect for those who did not have previous employment. SYEP enrollment could be interpreted more negatively if an individual was previously employed than if the individual was not, for example because it indicates that the individual was unable to secure re-employment with the previous employer.

the lottery with measures of peer group characteristics (including family income, gender, race, or age), the interactions are generally insignificant.

Appendix Tables (to be placed online)

Appendix Table 1: *Industry breakdown*

SYEP-Reported Job Type	Percent of sample	Imputed NAICS	Cluster
Arts and recreation	10.81	71	1
Camp (out of city)	10.59	72	1
Community/social service	11.06	62	1
Cultural institution	1.24	71	2
Day care/day camp	36.99	71	1
Educational services	7.68	61	1
Financial services	0.21	52	2
Government agency	7.02	92	1
Healthcare/medical	7.74	62	1
Hospitality/tourism	0.09	71	1
Legal services	0.20	54	2
Other	3.80	99	2
Real estate/property	1.14	53	2
Retail	1.24	44	2
Science and technology	0.19	54	2

Notes: The table shows the percentage of SYEP participants in each industry, as classified by SYEP administrative records. The "Cluster" column shows whether we classify a SYEP-reported industry designation into North American Industrial Classification System (NAICS) codes 61, 62, 71, and 92 (comprising Cluster 1), or into other NAICS codes (comprising Cluster 2). Cluster 1 consists of industries that are overrepresented in SYEP jobs relative to the control group's industry distribution.

Appendix Table 2: *Comparison of Compliers to Never-Takers*

(1) Dependent variable	(2) Coefficient (SE) on SYEP participation
A) Total earnings in Year -1	-233.43 (41.06)***
B) Job in Year -1	-0.012 (0.010)
C) College in Year -1	-0.026 (0.0022)***
D) Male	0.0055 (0.0064)
E) White	0.000044 (0.0047)
F) Black	0.045 (0.0076)***
G) Latino	-0.026 (0.0054)
H) Other race	-0.019 (0.0042)
I) Age	-0.42 (0.019)***
J) U.S. citizen dummy	0.025 (0.0027)***

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from OLS regressions of the indicated outcomes on the SYEP participation dummy and provider-by-year fixed effects. The sample is limited to those who won the SYEP lottery, so that the coefficient reflects the difference between compliers and never-takers. We show these results in regression form, rather than comparing raw means between compliers and never-takers, because the first stage differs across provider-years; thus, we need to control for provider-by-year fixed effects to eliminate cross-sectional variation across provider-year combinations. "Job in Year -1" refers to a dummy for having positive earnings in Year -1. The sample size in all regressions is 164,977. Age is expressed in years. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 3: *"Reduced form" effect of winning SYEP lottery on employment outcomes*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy
A) Year 0	639.63	792.25	-152.47	0.52	-0.035
	(22.31)***	(12.80)***	(17.81)***	(0.0097)***	(0.0026)***
B) Year 1	-72.71	33.56	-106.27	0.0090	-0.013
	(29.24)***	(3.78)***	(29.18)***	(0.0025)***	(0.0019)***
C) Year 2	-68.73	17.00	-85.73	0.0033	-0.0071
	(30.76)***	(2.52)***	(31.02)***	(0.0023)	(0.0020)***
D) Year 3	-80.89	6.00	-86.89	-0.00046	-0.0038
	(32.41)***	(1.55)***	(32.28)***	(0.0017)	(0.0017)**
E) Year 4	-25.66	3.26	-28.92	0.00097	-0.00024
	(32.89)	(1.17)***	(33.19)	(0.0016)	(0.0016)
F) Years 0-4	640.61	1167.46	-526.64	0.065	-0.0043
	(100.67)***	(14.83)***	(100.32)***	(0.0025)***	(0.0016)***
G) Years 1-4	-263.05	82.21	-345.26	0.0073	-0.0020
	(97.59)***	(9.38)***	(98.98)***	(0.0016)***	(0.0016)

Notes: The table shows the "intent-to-treat" estimates: coefficients and standard errors on the treatment dummy from OLS regressions of earnings and employment outcomes on a dummy for winning the SYEP lottery. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level. See other notes to Table 3.

Appendix Table 4: *Effect of SYEP participation on earnings and employment outcomes in Years 5 to 7*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy	Number of obs.	Number of individuals
Panel A: Without controls							
A) Year 5	9.93	0.85	9.08	0.0014	-0.00038	200,711	147,577
	(52.05)	(1.18)	(52.02)	(0.0024)	(0.0024)		
B) Year 6	-70.77	-0.23	-75.84	-0.0020	-0.00072	116,079	94,662
	(95.44)	(1.48)	(105.50)	(0.0033)	(0.0033)		
C) Year 7	271.19	-1.70	272.88	-0.0035	-0.0027	56,557	56,557
	(137.74)**	(1.97)	(137.95)**	(0.0051)	(0.0050)		
Panel B: With controls							
D) Year 5	-3.96	0.92	-4.88	0.00085	-0.00098	200,711	147,577
	(46.41)	(1.23)	(46.36)	(0.0025)	(0.0026)		
E) Year 6	-74.12	-0.26	-70.52	-0.0025	-0.0012	116,079	94,662
	(95.99)	(1.48)	(95.62)	(0.0033)	(0.0034)		
F) Year 7	167.10	-1.64	168.74	-0.0044	-0.0036	56,557	56,557
	(129.29)	(1.96)	(129.45)	(0.0052)	(0.0050)		

Notes: The table shows coefficients and standard errors on the treatment dummy from IV regressions of earnings and employment outcomes on SYEP participation, controlling for covariates. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. The controls added in Panel B are: gender, dummies for race categories, citizenship, age, number of family members, individual's wage income in Year -1, individual's NYC government wages in Year -1, individual's non-NYC government wages in Year -1, number of jobs in Year -1, a dummy for whether the individual had any job in Year -1, a dummy for whether the individual was enrolled in college in Year -1, a dummy for whether the individual was claimed on a return in Year -1, and family income in Year -1 if the individual was claimed. See other notes to Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 5: *Effect of SYEP participation on earnings and employment outcomes with controls.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy
A) Year 0	903.65	1085.25	-181.38	0.71	-0.047
	(16.26)***	(10.11)***	(12.92)***	(0.0065)***	(0.0032)***
B) Year 1	-74.83	46.14	-120.97	0.013	-0.017
	(23.51)***	(4.90)***	(24.09)***	(0.0033)***	(0.0027)***
C) Year 2	-72.63	23.38	-96.01	0.0044	-0.0098
	(28.25)***	(3.34)***	(29.02)***	(0.0029)	(0.0024)***
D) Year 3	-92.94	8.23	-101.19	-0.00090	-0.0054
	(31.57)***	(2.13)***	(31.43)***	(0.0021)	(0.0022)**
E) Year 4	-22.70	4.46	-27.16	0.00093	-0.00073
	(36.76)	(1.57)***	(37.13)	(0.0022)	(0.0021)
F) Years 0-4	640.61	1167.46	-526.64	0.088	-0.0062
	(100.67)***	(14.83)***	(100.32)***	(0.0033)***	(0.0021)***
G) Years 1-4	-263.05	82.21	-345.26	0.0098	-0.0030
	(97.59)***	(9.38)***	(98.98)***	(0.0021)***	(0.0021)

Notes: The table shows coefficients and standard errors on the treatment dummy from IV regressions of earnings and employment outcomes on SYEP participation, controlling for covariates. The instrument for whether an individual participated in SYEP is a dummy indicating whether an individual won the SYEP lottery. The table is based on the specification as Table 3 in the main text, except that we control for the covariates listed in Appendix Table 4. Controlling for any subset of these covariates yields extremely similar results. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 6: *Effect of SYEP participation on earnings and employment outcomes, robustness tests.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy
Panel A. Initial lottery					
A) Year 0	869.61	1116.22	-246.48	0.70	-0.057
	(37.58)***	(9.28)***	(37.82)***	(0.0064)***	(0.0040)***
B) Years	-448.31	81.44	-529.76	0.011	-0.00052
1-4	(170.92)***	(12.33)***	(169.44)***	(0.0024)***	(0.0025)
Panel B. Match only on SSN					
C) Year 0	876.96	1085.60	-208.42	0.71	-0.048
	(25.17)***	(10.15)***	(24.89)***	(0.0063)***	(0.0035)***
D) Years	-336.85	81.90	-418.77	0.010	-0.0026
1-4	(155.60)**	(9.53)***	(156.00)***	(0.0021)***	(0.0021)
Panel C. Cluster by individual					
E) Year 0	876.26	1085.34	-208.87	0.71	-0.048
	(26.62)***	(10.44)***	(26.70)***	(0.0021)***	(0.0024)***
F) Years	-339.73	81.96	-421.70	.010	-0.0027
1-4	(142.93)**	(5.31)***	(143.54)***	(0.0017)***	(0.0021)

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from IV regressions of employment outcomes on SYEP participation. In Panel A, the instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the *initial* SYEP lottery. In Panel B, the regressions are identical to those in the baseline specification in Table 3, except that we include people only if their SSN matches between the SYEP and IRS data. The sample size in Panel B is 293,908. In Panel C, we cluster the standard errors at the level of the individual. In all of these cases, the results are similar to the main results in Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 7. *Effect of SYEP participation on earnings and employment outcomes, dynamic specification.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy
A) Year 0	876.26	1085.34	-208.87	0.71	-0.048
	(25.28)***	(10.22)***	(25.03)***	(0.0065)***	(0.0042)***
B) Year 1	-126.40	12.79	-139.20	-0.012	-0.024
	(39.84)***	(2.29)***	(39.71)***	(0.0052)***	(0.0051)***
C) Year 2	-104.93	9.95	-114.88	-0.0038	-0.013
	(42.46)**	(2.08)***	(42.49)***	(0.0057)	(0.0057)**
D) Year 3	-110.61	4.19	-114.81	-0.0015	-0.0048
	(43.65)**	(1.39)***	(43.53)***	(0.0058)	(0.0057)
E) Year 4	-31.96	2.51	-34.47	-0.00077	-0.0035
	(43.81)	(1.32)*	(44.18)	(0.0061)	(0.0060)
F) Years 0-4	502.36	1114.79	-612.24	--	--
	(170.95)***	(11.54)***	(171.24)***		
G) Years 1-4	-373.90	29.44	-403.37	--	--
	(152.24)**	(4.52)***	(152.33)***		

Notes: This table employs the dynamic IV estimator of Cellini, Ferreira, and Rothstein (2010), as described in the text. The table shows coefficients and standard errors on the SYEP participation dummy from IV regressions of employment outcomes on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. We perform the dynamic estimate of the effect on earnings in Years 0-4 (or 1-4) by summing the coefficients estimated in the dynamic specification from Years 0 to 4 (or 1-4, respectively). "--" indicates that we do not perform the estimates from Years 0-4 and 1-4 for the probability of having a job; we cannot add these coefficients across Rows A through E (as in the case of the earnings estimates) because the probabilities are not independent. See other notes to Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 8. *Effect of SYEP participation on earnings outcomes by SYEP participation history.*

	(1) Rule out prior SYEP participation (SYEP data)	(2) No prior SYEP participation (IRS data)	(3) Prev. participated once (IRS data)	(4) IV for number of times participated	(5) Total earnings of other family members
A) Year 0	701.33	651.88	671.18	948.49	-41.75
	(37.83)***	(28.46)***	(24.79)***	(27.82)***	(187.10)
B) Year 1	-50.74	-54.87	-125.56	-103.28	33.57
	(38.30)***	(32.22)*	(38.37)***	(41.54)***	(180.08)
C) Year 2	-90.18	-69.50	-78.79	-95.60	1.61
	(40.45)**	(33.85)***	(54.92)	(42.75)**	(175.93)
D) Year 3	-98.76	-79.88	-76.49	-111.75	-10.72
	(40.46)**	(36.64)***	(80.74)	(44.87)**	(175.09)
E) Year 4	-70.58	-18.20	-41.19	-35.29	-32.78
	(47.92)	(43.29)	(93.05)	(44.96)	(182.58)
N	115,337	202,717	60,437	294,580	294,580

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from IV regressions of earnings outcomes on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. Column 1 shows results for applicants who were young enough that they never could have been eligible to participate in SYEP previously, because they were 14 or younger in 2005. In Columns 2 and 3, we use IRS records on NYC government earnings in prior years as a proxy for prior SYEP participation (classifying individuals as having participated in SYEP in a given prior year 1999 or after when they had positive NYC government earnings in that year). We show results for those who had previously participated no times or one time in Columns 2 and 3, respectively. For those participating two or more times, the sample sizes are much smaller, and the results are insignificant and uninformative. In Column 4, we use winning the SYEP lottery as an instrument for the total number of times participating in SYEP between 1999 and Year 0 (inclusive). In Column 5, we examine the effect on the total earnings of other family members. See other notes to Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 9. *Effect of SYEP participation on earnings outcomes by quantile.*

	(1) Year 0	(2) Year 1	(3) Year 2	(4) Year 3	(5) Year 4
Panel A: Median regressions					
A) Total	1101.24	83.40	23.91	-41.87	-54.99
	(2.77)***	(18.83)***	(14.51)*	(22.09)*	(31.28)
B) NYC gov't	1077.87	--	--	--	--
	(1.45)***				
C) Non-NYC gov't	--	--	-54.50	-46.32	6.73
			(17.40)***	(24.58)*	(34.11)
Panel B: Regressions for 75th percentile					
D) Total	477.80	-134.90	-122.13	-140.45	7.59
	(2.72)***	(27.69)***	(39.96)***	(52.50)***	(60.44)
E) NYC gov't	1181.88	--	--	--	--
	(0.98)***				
F) Non-NYC gov't	-129.72	-177.46	-155.49	-145.19	2.42
	(11.03)***	(28.29)***	(44.33)***	(52.43)***	(61.04)
Panel C: Regressions for 90th percentile					
G) Total	88.76	-254.62	-188.26	-204.80	-9.35
	(37.01)**	(64.94)***	(71.74)***	(84.75)**	(105.87)
H) NYC gov't	1088.80	16.04	71.18	--	--
	(1.21)***	(1.02)***	(8.25)***		
I) Non-NYC gov't	-571.83	-288.38	-201.47	-365.84	-27.47
	(37.97)***	(66.53)***	(72.40)***	(87.35)***	(107.31)
Panel D: Descriptive statistics					
J) Percent earning 0	39.88	44.13	38.77	33.24	28.83
K) Median total	939.18	629.75	1,051.60	1,386.65	2,473.01
L) 75th total	1,353.58	2,006.57	3,947.14	6,196.07	8,674.60
M) 90th total	2,984.16	6,293.88	9,613.63	12,852.71	16,378.32

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from quantile regressions of earnings outcomes on SYEP participation. The independent variable of interest is a dummy indicating that an individual won the SYEP lottery. Each row investigates a different quantile and/or outcome variable. Panel A investigates the effect of SYEP on median earnings; Panel B investigates the effect on the 75th percentile of earnings; and Panel C investigates the effect on the 90th percentile. Rows A, D, and G investigate the effect on total earnings; B, E, and H investigate the effect on NYC government earnings; and C, F, and I investigate the effect on non-NYC government earnings. For context, Panel D shows descriptive statistics. "--" indicates that the quantile of earnings in question is zero, which implies that SYEP participation has no effect on the quantile in question. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level. See other notes to Table 3.

Appendix Table 10a. *Effect of SYEP participation on Year 0 earnings outcomes among subsamples.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) *p*-value for Total Earnings	(5) N
A) WOTC-eligible	842.21	1089.02	-246.66	0.58	32,248
	(63.01)***	(11.46)***	(62.90)***		
B) WOTC-ineligible	879.04	1084.97	-205.71		262,332
	(26.40)***	(10.15)***	(26.30)***		
C) Below-median inc.	888.81	1086.79	-197.91	0.68	147,291
	(27.58)***	(13.06)***	(22.34)***		
D) Above-median inc.	860.00	1083.95	-223.60		147,289
	(53.92)***	(7.99)***	(54.21)***		
E) Males	855.87	1079.51	-223.53	0.48	132,692
	(31.98)***	(9.74)***	(32.20)***		
F) Females	892.25	1090.20	-197.65		161,888
	(38.69)***	(10.76)***	(39.03)***		
G) White	899.28	1,195.37	-296.09	0.37	37,172
	(70.66)***	(42.28)***	(72.58)***		
H) Black	899.14	1,073.35	-174.02		142,627
	(34.88)***	(5.88)***	(33.68)***		
I) Latino	803.95	1,055.55	-251.37		79,095
	(53.72)***	(8.59)***	(52.00)***		
J) Other races	944.36	1,134.86	-190.96		35,686
	(65.63)***	(34.09)***	(57.07)***		
K) Older	688.58	1068.93	-380.13	0.00000	147,260
	(41.55)***	(12.69)***	(38.86)***		
L) Younger	1033.38	1099.51	-65.91		147,320
	(27.50)***	(8.84)***	(28.03)**		
M) Work in Year -1	691.98	1090.66	-397.98	0.00090	94,855
	(80.32)***	(15.00)***	(82.92)***		
N) No work in Year -1	966.08	1083.04	-116.96		199,725
	(12.40)***	(9.07)***	(8.82)***		
O) 2005-6 lotteries	903.03	1068.96	-165.84	0.39	116,079
	(34.51)***	(11.08)***	(35.24)***		
P) 2007-8 lotteries	858.58	1096.16	-237.28		178,501
	(36.17)***	(10.62)***	(35.07)***		
Q) Emp. Zones	886.86	1082.32	-195.31	0.69	149,379
	(34.35)***	(10.50)***	(34.21)***		
R) Non-Emp. Zones	866.47	1088.42	-221.66		145,200
	(37.00)***	(10.90)***	(36.47)***		

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from IV regressions of employment outcomes on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. Each row shows the results for a different population. "Work in Year -1" indicates that the individual had positive earned income in Year -1. Below-median income refers to individuals with family income in Year -1 of $26,313 and below, and above-median income refers to individuals in families with higher income (where the median income refers to the median in Year -1 in the sample we investigate). The Older category is at least age 16.25, whereas the Younger category is below this age. The sample size is slightly different for above-median and below-median incomes, and for older and younger ages, because multiple individuals have the median value of income and age. Column 4 shows the *p*-value of the test of equality of the coefficients across the groups in question, when the dependent variable is total earnings in Year 0 (the tests of equality for non-NYC government earnings show similar results). See other notes to Table 3. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 10b. *Effect of SYEP participation on Years 1-4 earnings outcomes among subsamples.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) *p*-value for Total Earnings	(5) N
A) WOTC-eligible	-18.73	57.84	-72.56	0.41	32,248
	(429.52)	(18.65)***	(429.80)		
B) WOTC-ineligible	-384.76	84.98	-469.75		262,332
	(159.92)**	(9.90)***	(160.54)***		
C) Below-median inc.	-317.61	74.59	-392.20	0.82	147,291
	(148.55)**	(11.35)***	(150.49)***		
D) Above-median inc.	-382.17	89.24	-471.43		147,289
	(256.07)	(10.66)***	(256.41)*		
E) Males	-365.47	92.29	-457.78	0.88	132,692
	(162.55)**	(11.61)***	(166.49)***		
F) Females	-321.49	72.92	-394.41		161,888
	(243.72)	(10.46)***	(244.62)		
G) White	-1,242.35	120.60	-1,362.95	0.27	37,172
	(549.18)**	(20.56)***	(554.96)**		
H) Black	-122.69	72.27	-194.93		142,627
	(201.26)	(13.93)***	(203.38)		
I) Latino	-520.60	92.48	-613.17		79,095
	(329.05)	(10.71)***	(327.03)*		
J) Other races	-209.49	71.49	-280.98		35,686
	(296.54)	(15.19)***	(300.37)		
K) Older	-726.62	74.33	-800.96	0.013	147,260
	(250.08)***	(9.36)***	(251.94)***		
L) Younger	-44.06	88.33	-132.41		147,320
	(158.11)	(11.35)***	(155.13)		
M) Work in Year -1	-993.54	56.70	-1050.28	0.021	94,855
	(419.67)**	(11.84)***	(422.72)**		
N) No work in Year -1	5.02	92.50	-87.48		199,725
	(101.75)	(10.79)***	(102.42)		
O) 2005-6 lotteries	-61.98	77.44	-139.45	0.13	116,079
	(235.76)	(13.66)***	(238.65)		
P) 2007-8 lotteries	-523.10	84.94	-608.04		178,501
	(195.88)***	(9.82)***	(196.48)***		
Q) Emp. Zones	-299.07	80.58	-379.69	0.82	149,379
	(193.43)	(8.23)***	(193.47)**		
R) Not Emp. Zones	-363.00	83.20	-446.18		145,200
	(222.90)	(13.79)***	(222.38)**		

Notes: The table is identical to Table 10a above, except that the dependent variable in the regressions in 10b is earnings in Years 1-4. See other notes to Table 10a.

Appendix Table 11: *Effect of SYEP participation on job transitions.*

	(1) Coefficient (standard error)	(2) Mean of dependent variable
A) Year 0	-0.029	0.47
	(0.0084)***	
B) Year 1	-0.012	0.21
	(0.0067)*	
C) Year 2	-0.0037	0.13
	(0.0054)	
D) Year 3	-0.0033	0.09
	(0.0040)	
E) Year 4	0.0038	0.06
	(0.0034)	

Notes: The table shows coefficients and standard errors on a dummy for participating in SYEP, from a two-stage least squares regression. The instrument for participating in SYEP is whether an individual won the SYEP lottery. The dependent variable is the fraction of employers that an individual worked at in Year -1 that the individual still worked at in a given year. All regressions control for provider-year dummies. The second column shows the mean of the dependent variable. All regressions have 38,808 observations; the sample size is smaller than the main sample because the sample is limited to individuals who had a non-SYEP job in Year -1. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 12: *Effect of SYEP provider industry mix on earnings outcomes.*

	(1) Total Earnings	(2) Total Earnings in Cluster 1	(3) Total Earnings in Cluster 2	(4) Total NYC gov't earnings	(5) Total non-NYC gov't earnings
Panel A: interaction term					
A) Year 0	-407.30	905.01	-1325.21	118.94	-526.17
	(219.90)*	(98.65)***	(172.97)***	(124.66)	(201.77)***
B) Year 1	-473.15	-40.71	-439.82	-43.44	-429.73
	(304.07)	(105.07)	(245.44)*	(48.50)	(288.15)
C) Year 2	-483.10	-108.85	-378.96	4.84	-487.94
	(306.31)	(99.45)	(271.40)	(34.73)	(293.13)
D) Year 3	-622.77	-28.06	-604.60	-17.06	-605.57
	(240.40)***	(158.76)	(232.93)***	(24.33)	(238.78)***
E) Year 4	-643.19	61.95	-688.00	-13.72	-629.47
	(296.82)**	(177.52)	(230.35)***	(11.53)	(298.94)**
F) Years	-2629.51	789.34	-3436.59	49.56	-2678.88
0-4	(1086.95)**	(472.85)	(937.57)***	(139.15)	(1106.98)**
G) Years	-2222.21	-115.66	-2111.38	-69.38	-2152.71
1-4	(961.25)**	(447.10)	(811.91)**	(91.13)	(945.14)**
Panel B: main effect					
A) Year 0	1011.71	-126.13	1150.93	680.23	331.57
	(202.92)***	(90.74)	(157.82)***	(115.26)***	(186.71)*
B) Year 1	365.40	60.33	310.13	76.14	289.28
	(281.69)	(96.18)	(225.75)	(45.22)*	(267.83)
C) Year 2	376.07	112.30	267.88	12.49	363.58
	(282.07)	(88.53)	(250.90)	(32.78)	(269.62)
D) Year 3	491.07	19.55	481.37	21.57	469.37
	(219.59)**	(143.80)	(218.62)**	(22.37)	(217.80)**
E) Year 4	566.75	-73.60	624.16	15.90	550.86
	(269.55)**	(160.16)	(212.78)***	(10.77)	(271.33)**
F) Years	2811.01	-7.55	2834.46	806.32	2004.66
0-4	(1005.44)***	(426.76)	(875.69)***	(131.53)***	(1025.63)*
G) Years	1799.30	118.59	1683.53	126.09	1673.09
1-4	(888.05)**	(401.58)	(760.78)**	(85.45)	(873.25)*

Notes: The table shows the results of OLS regressions in which the dependent variable is earnings (where the particular type of earnings in question is shown in the column heading). The independent variables are: 1) a variable formed by interacting a dummy for winning the SYEP lottery with the percent of the provider that is in Industry Cluster 1; 2) a dummy for winning the SYEP lottery; and 3) dummies for each provider-lottery combination. Panel A shows coefficients on the variable formed by interacting a dummy for winning the SYEP lottery with the percent of the provider that is in Industry Cluster 1. Panel B shows coefficients and standard errors on the dummy for winning the SYEP lottery. See notes to Tables 3 and 4. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 13. *Effect of SYEP participation on college attendance, with controls*

	Coefficient (SE) on SYEP participation
A) Year 0	0.00035
	(0.0012)
B) Year 1	0.0024
	(0.0015)
C) Year 2	-0.0016
	(0.0025)
D) Year 3	0.00010
	(0.0022)
E) Year 4	-0.0036
	(0.0023)
F) Years 0-4	-0.0042
	(0.0024)*
G) Years 1-4	-0.0035
	(0.0024)
H) Total years of college	-0.0023
	(0.0065)

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions of a college attendance dummy or total years of college on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. The table is identical to Table 5, except that we add controls for the demographics listed in Appendix Table 4. See other notes to Table 5. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 14. *Effect of SYEP participation on earnings and employment outcomes for those 18 years and older.*

	(1) Total Earnings	(2) NYC gov't earnings	(3) Non-NYC gov't earnings	(4) Job dummy	(5) Non-NYC gov't job dummy
A) Year 0	535.44	1039.67	-504.02	0.40	-0.066
	(73.63)***	(15.83)***	(71.65)***	(0.013)***	(0.0065)**
B) Year 1	-329.11	45.42	-374.53	0.0016	-0.016
	(122.02)**	(5.09)***	(123.34)***	(0.0064)	(0.0076)**
C) Year 2	-287.46	20.27	-307.72	-0.0012	-0.0072
	(132.09)**	(4.12)***	(133.17)**	(0.0064)	(0.0066)
D) Year 3	-311.43	10.10	-321.53	0.000037	-0.0047
	(158.26)**	(4.18)***	(159.04)**	(0.0055)	(0.0056)
E) Year 4	-139.65	8.38	-148.03	0.0029	0.0011
	(172.48)	(3.23)***	(173.76)	(0.0049)	(0.0050)
F) Years 0-4	-532.20	1123.83	-1655.82	0.059	-0.0056
	(588.31)	(20.28)***	(591.56)	(0.0063)***	(0.0036)
G) Years 1-4	-1067.64	84.16	-1151.80	0.0063	-0.0014
	(528.85)**	(12.05)***	(533.68)**	(0.0038)*	(0.0041)

Notes: The table shows coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions of earnings and employment outcomes on SYEP participation. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. The sample is limited to those 18 years of age or older in the year of SYEP participation, but the table is otherwise based on the identical specification to Table 3. The sample size is 72,432.

Appendix Table 15. *Effect of SYEP participation on incarceration: alternative specifications.*

	(1) 2SLS, with covariates	(2) Times incarcerated	(3) Probit	(4) Prison dummy mean (x 100)
(A) Full population	-0.092 (0.045)**	-0.11 (0.049)**	-0.064 (0.029)**	0.95
(B) 19 and over	-0.49 (0.22)**	-0.44 (0.25)*	-0.39 (0.18)**	0.85
(C) 18 and under	-0.065 (0.045)	-0.085 (0.051)*	-0.047 (0.030)	0.96
D) Below-median inc.	-0.14 (0.074)*	-0.13 (0.077)*	-0.093 (0.049)*	1.14
E) Above-median inc.	-0.051 (0.064)	-0.084 (0.073)	-0.039 (0.042)	0.76
(F) Males	-0.22 (0.094)**	-0.24 (0.098)**	-0.16 (0.066)**	2.01
(G) Females	0.023 (0.020)	0.015 (0.021)	0.027 (0.024)	0.08
(H) White	-0.15 (0.086)*	-0.16 (0.096)*	-0.12 (0.050)***	0.07
(I) Black	-0.16 (0.070)**	-0.19 (0.077)***	-0.13 (0.052)**	1.48
(J) Latino	0.051 (0.078)	0.041 (0.085)	0.027 (0.056)	0.70
(K) Other	-0.062 (0.098)	-0.038 (0.11)	-0.16 (0.24)	0.33
(N) Prior work	-0.037 (0.091)	-0.055 (0.092)	-0.034 (0.056)	0.79
(O) No prior work	-0.12 (0.049)**	-0.14 (0.053)***	-0.082 (0.033)**	1.02
P) 2005-6 lotteries	-0.11 (0.078)	-0.15 (0.086)*	-0.088 (0.057)	1.15
Q) 2007-8 lotteries	-0.084 (0.056)	-0.081 (0.063)	-0.052 (0.032)	0.82

Notes: Columns 1 and 2 show coefficients and standard errors on the SYEP participation dummy from two-stage least squares regressions of a dummy for incarceration in NYS on SYEP participation. In Column 3, the dependent variable is the number of times incarcerated. The instrument for whether an individual participated in SYEP is a dummy indicating that an individual won the SYEP lottery. Column 4 shows coefficients and standard errors from a probit regression of the incarceration dummy on the SYEP participation dummy. Before running the regressions, we multiply the dependent variable by 100 so that coefficients show percentage point changes (for the reader's ease). Note that the probit specification runs "reduced form" regressions that regress the dependent variable directly on the dummy for winning the SYEP lottery, not an instrumental variables regression. The probit coefficients represent marginal effects, calculated at the mean. Adding controls to the specifications in Columns 2 and 3 yields nearly identical results. In addition to the groups discussed in the main text, the table also shows the robustness checks for additional groups discussed in Gelber, Isen, and Kessler (2015). See Table 6 and Appendix Table 10a for other notes and information on samples. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 16. *Effect of SYEP participation on incarceration among those 19 and older at the time of SYEP participation*

	(1) 2SLS	(2) Prison dummy mean (x 100)	(3) N
A) Below-median inc.	-0.69 (0.31)**	0.96	12,405
B) Above-median inc.	-0.30 (0.28)	0.74	12,404
(C) Males	-0.90 (0.50)*	1.77	10,633
(D) Females	-0.14 (0.0012)	0.16	14,176
(E) White	-0.28 (0.20)	0.064	4,652
(F) Black	-0.20 (0.40)	1.44	11,010
(G) Latino	-1.12 (0.33)***	0.71	5,200
(H) Other	-0.42 (0.42)	0.30	3,947
(I) Prior work	-0.25 (0.23)	0.72	16,083
(J) No prior work	-0.88 (0.045)**	1.09	8,726
K) 2005-6 lotteries	-0.40 (0.35)	0.76	8,119
L) 2007-8 lotteries	-0.53 (0.28)*	0.89	16,690
M) Emp. Zone	-0.66 (0.25)***	1.32	12,719
N) Non-Emp. Zone	-0.30 (0.35)	0.85	12,090

Notes: This table estimates the effect of SYEP participation on incarceration for those 19 or over at the time of SYEP participation. Column 3 shows the sample size in each group. See Appendix Table 15 for other notes. Table 6 row B also shows the effect for the full 19 and older group. The results are nearly identical under other specifications shown in Appendix Table 15. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 17. *Effect of SYEP on mortality by calendar year.*

	(1) 2SLS	(2) 2SLS, with controls	(3) Mortality dummy mean (x 100)
A) 2005	0.0075	0.0076	0.011
	(0.0091)	(0.0091)	
B) 2006	-0.0041	0.0042	0.021
	(0.011)	(0.011)	
C) 2007	-0.00069	-0.00050	0.035
	(0.011)	(0.011)	
D) 2008	-0.018	-0.017	0.055
	(0.011)*	(0.010)*	
E) 2009	-0.0043	-0.0037	0.10
	(0.018)	(0.017)	
F) 2010	-0.014	-0.013	0.16
	(0.021)	(0.021)	
G) 2011	-0.033	-0.032	0.21
	(0.025)	(0.025)	
H) 2012	-0.043	-0.041	0.27
	(0.024)*	(0.024)*	
I) 2013	-0.061	-0.059	0.34
	(0.028)**	(0.028)**	
J) 2014	-0.075	-0.073	0.38
	(0.031)**	(0.031)**	

Notes: The table shows estimates of the effect of SYEP participation on mortality using a two-stage least squares, linear probability model. Each row shows the results for a different calendar year. We show the effect of SYEP on a dummy for whether an applicant died *by* a given year; thus, the effect *in* a given year can be calculated as the difference between the coefficient for that year and the previous year. Column 1 shows the results of our two-stage least squares specification (1)-(2). Column 2 shows the results of this specification when we add the controls listed in Appendix Table 4. Column 3 shows the mean of the dependent variable (i.e. the dummy measuring the probability of mortality by each year, relative to year of SYEP participation). So that readers can more easily interpret the results, we have multiplied the mortality dummy by 100. The results are comparable with Cox or probit models; we show a two-stage least squares (linear probability) model here to show results that are comparable to the IV results elsewhere in the paper. We use data through October 2014. See Table 7 for other notes and information on samples. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 18. *Effect of SYEP on mortality by year since lottery.*

	(1) 2SLS	(2) 2SLS, with controls	(3) Mortality dummy mean (x 100)
A) Year 0	0.0021	0.0021	0.01
	(0.0050)	(0.0050)	
B) Year 1	-0.000083	0.00013	0.04
	(0.011)	(0.011)	
C) Year 2	-0.0044	-0.0038	0.08
	(0.016)	(0.016)	
D) Year 3	-0.013	-0.012	0.14
	(0.021)	(0.021)	
E) Year 4	-0.015	-0.014	0.19
	(0.022)	(0.022)	
F) Year 5	-0.027	-0.025	0.25
	(0.025)	(0.025)	
G) Year 6	-0.028	-0.026	0.30
	(0.029)	(0.029)	
H) Year 7	-0.099	-0.096	0.38
	(0.036)***	(0.035)***	
I) Year 8	-0.15	-0.14	0.44
	(0.053)***	(0.052)***	
J) Year 9	-0.22	-0.22	0.55
	(0.085)***	(0.085)***	

Notes: The table shows estimates of the effect of SYEP participation on mortality using a two-stage least squares, linear probability model. Each row shows the results for a different year relative to the year of SYEP participation. We show the effect of SYEP on a dummy for whether an applicant died *by* a given year; thus, the effect *in* a given year can be calculated as the difference between the coefficient for that year and the previous year. Column 1 shows the results of our two-stage least squares specification (1)-(2). Column 2 shows the results of this specification when we add the controls listed in Appendix Table 4. Column 3 shows the mean of the dependent variable (i.e. the dummy measuring the probability of mortality by each year, relative to year of SYEP participation). So that readers can more easily interpret the results, we have multiplied the mortality dummy by 100. The results are comparable with Cox or probit models; we show a two-stage least squares (linear probability) model here to show results that are comparable to the IV results elsewhere in the paper. See Table 7 for other notes and information on samples. Because the data extend until 2014, we observe lotteries from all four years (2005, 2006, 2007, and 2008) only until Year 6, implying that sample sizes are not constant across Years 6 to 9; see Appendix Table 4 for sample sizes in the relevant sets of lotteries. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 19. *Effect of SYEP on mortality.*

	(1) 2SLS, with covariates	(2) Cox	(3) Probit	(4) Mortality dummy mean (x 100)
A) Full population	-0.073	0.86	-0.055	0.38
	(0.031)**	(0.054)**	(0.023)**	
B) WOTC-eligible	-0.10	0.82	-0.074	0.38
	(0.10)	(0.16)	(0.074)	
C) WOTC-ineligible	-0.070	0.87	-0.052	0.38
	(0.030)**	(.053)**	(0.022)**	
D) Below-median inc.	-0.086	0.86	-0.063	0.42
	(0.061)	(0.093)	(0.044)	
E) Above-median inc.	-0.060	0.87	-0.046	0.34
	(0.037)	(0.072)*	(0.028)*	
F) Males	-0.15	0.83	-0.11	0.61
	(0.059)**	(0.062)**	(0.043)**	
G) Females	-0.015	0.94	-0.012	0.19
	(0.032)	(0.12)	(0.023)	
H) White	0.0044	1.01	0.0018	0.16
	(0.076)	(0.32)	(0.047)	
I) Black	-0.058	0.91	-0.047	0.52
	(0.048)	(0.068)	(0.037)	
J) Latino	-0.14	0.72	-0.10	0.32
	(0.056)**	(0.092)**	(0.040)**	
K) Other races	-0.052	0.80	-0.038	0.18
	(0.073)	(0.22)	(0.052)	
L) Older	-0.027	0.95	-0.020	0.40
	(0.047)	(0.079)	(0.033)	
M) Younger	-0.11	0.78	-0.087	0.33
	(0.040)***	(0.070)***	(0.031)***	
N) Work in Year -1	0.041	1.07	0.024	0.35
	(0.059)	(0.13)	(0.042)	
O) No work in Year -1	-0.11	0.79	-0.092	0.39
	(0.043)***	(0.065)***	(0.032)***	
P) 2005-6 lotteries	-0.16	0.76	-0.13	0.47
	(0.051)***	(0.063)***	(0.040)***	
Q) 2007-8 lotteries	-0.015	0.96	-0.012	0.32
	(0.044)	(0.092)	(0.030)	

Notes: Column 1 shows hazard ratios and standard errors on a dummy for winning the SYEP lottery from a right-censored Cox proportional hazard model of time to mortality. Column 2 shows coefficients and standard errors from a probit regression. Each row shows the results for a different population. We eliminate from the regressions those rare cases of individuals who died between applying to SYEP and the date of first participating in SYEP. See Table 7 and Appendix Table 10 for other notes and information on samples. The Cox and probit specifications run "reduced form" regressions that regress the dependent variable directly on the dummy for winning the SYEP lottery, not an instrumental variables regression. The probit coefficients represent marginal effects, calculated at the mean. Column 3 shows the mean of the "mortality by 2014" dummy, multiplied by 100, in each group. So that readers can more easily interpret the results, we have also multiplied the dependent variable by 100 in Columns 1 and 2. In addition to the groups discussed in the main text, the table also shows the robustness checks for additional groups discussed in Gelber, Isen, and Kessler (2015). *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.

Appendix Table 20. *Effect of SYEP on key outcomes with different first stage for each provider.*

(1) Dependent variable	(2) Coefficient (SE) on SYEP participation with separate first stage for each provider	(3) Coefficient (SE) on SYEP participation with separate first stage for each provider-year
A) Year 0 Total Earnings	878.06 (26.30)***	877.52 (26.00)***
B) Year 0 Job	0.71 (0.0021)***	0.71 (0.0021)***
C) Years 1-4 Total Earnings	-330.22 (140.56)**	-319.79 (140.11)**
D) Years 1-4 Job	0.010 (0.0017)***	0.010 (0.0017)***
E) Years 0-4 Total Earnings	547.85 (159.65)***	557.73 (158.81)***
F) Years 0-4 Job	0.089 (0.0015)***	0.088 (0.0014)***
G) Total years of College	0.0013 (0.0077)	0.0018 (0.0076)
H) Incarceration by 2013	-0.099 (0.054)*	0.096 (0.048)**
I) Mortality by 2014	-0.076 (0.034)**	-0.0090 (0.035)**

Notes: The regressions run in Column 2 are the same as the corresponding regressions in Tables 3, 5, 6, and 7, except that in our first stage regression we allow for a different first stage for each provider, by interacting the lottery win dummy with the 62 provider dummies as excluded instruments. Similarly, the regressions run in Column 3 are the same as the corresponding regressions in Tables 3, 5, 6, and 7, except that in our first stage regression we allow for a different first stage in each provider-year combination, by interacting the lottery win dummy with the dummies for each provider-year combination as excluded instruments. For incarceration and mortality, we run the 2SLS linear specification with the binary dependent variable, where the dependent variable has been multiplied by 100 for ease of interpretation (as in previous tables). See other notes to Tables 3, 5, 6, and 7. *** denotes significance at the 1% level; ** at the 5% level; and * at the 10% level.