# Information Frictions and Skill Signaling in the Youth Labor Market*

Sara B. Heller (University of Michigan & NBER)†
Judd B. Kessler (University of Pennsylvania & NBER)

June 24, 2022

**Abstract**

This paper demonstrates that information frictions limit the labor market trajectories of young people in the U.S. We provide credible skill signals—recommendation letters based on supervisor feedback—to a random subset of 43,409 participants in New York City's summer jobs program. Letters increase employment the following year by 3 percentage points (4.5 percent). Earnings effects grow over 4 years to a cumulative $1,349 (4.9 percent). We find no evidence of increased job search or confidence; instead, the signals help employers better identify successful matches with high-productivity workers. But the additional work hampers on-time high school graduation, especially among low-achieving students.

*This work was funded by the Social Policy Research Initiative at J-PAL North America. We thank the New York City Department of Youth and Community Development, the New York State Department of Labor, and the New York City Department of Education for sharing data with us. Thanks to Charlie Brown, David Deming, Harry Holzer, Alicia Modestino, Mike Mueller-Smith, Alex Rees-Jones, Ana Reynoso, Basit Zafar, and the JIM group at Princeton for helpful comments. We are particularly grateful to Julia Breitman at DYCD for all of her help along the way, to Ben Cosman at DOE for all his support, and to Alex Hirsch, Ashley Litwin, and Lauren Shaw for phenomenal project management and research assistance. All views in the paper are those of the authors and do not represent the positions of any data provider or government agency.

†Corresponding author: sbheller@umich.edu

# 1 Introduction

The challenges that young people face in the labor market, and the bigger barriers facing Black and Hispanic youth, have been a central focus of both research and policy for over 50 years (Freeman and Wise 1982; Freeman and Holzer 1986; Heinrich and Holzer 2011; Hoynes, Miller, and Schaller 2012; Kahn 2010). Half a century of active labor market programs have spent billions of dollars trying to improve youths' labor market outcomes. Yet despite some success in U.S. sector-focused training, and more frequent success in developing countries, youth labor market programs in high-income countries quite frequently fail (Card, Kluve, and Weber 2018; Katz et al. 2020; Crépon and Van Den Berg 2016).

Theory offers one potential explanation for why it is so hard to improve labor market success among young people: information frictions may constrain youth employment, even when an applicant has the appropriate skills to succeed in a job (Altonji and Pierret 2001; Farber and Gibbons 1996; Jovanovic 1979). On the demand side, employers may have difficulty anticipating an applicant's future productivity. Combined with screening costs or the cost of hiring unqualified workers, such information frictions may leave qualified applicants unemployed, mismatched, or underpaid. In addition, if employers statistically discriminate based on age, class, or race, such information frictions could help to explain disparities among these groups. Information frictions may also be present on the supply side, if young people lack the networks, knowledge, or confidence to complete a successful job search (Gonzalez and Shi 2010; Holzer 1988).

A small set of experiments on skill signaling demonstrates that better information can improve labor market outcomes in an online marketplace (Pallais 2014; Stanton and Thomas 2016) and in the developing economies of South Africa (Abel, Burger, and Piraino 2020; Carranza et al. 2020), Ethiopia (Abebe et al. 2021), and Uganda (Bassi and Nansamba 2021). One feature of these environments is that they are relatively low-information settings where signals may be particularly scarce. On oDesk, employers have almost no verifiable information on applicants (Pallais 2014). And in the African context, high rates of youth unemployment, a high prevalence of self-employment, and the lack of clear educational signals all contribute to an environment where work and schooling histories provide relatively little information about an applicant's potential productivity or match quality (Bandiera et al. 2021; Van der Berg 2007).[1] As a result, we do not know whether information frictions

1. Donovan, Lu, and Schoellman (2020) provide evidence that it is precisely the low-information environment that explains labor market dynamics like elevated transition rates and higher turnover rates among low tenure workers in developing countries. Indeed, a more significant role for information frictions is one potential explanation for why active labor market programs seem to work better in developing countries than in the U.S. or Europe.

constrain youth employment in high-income countries, where observable characteristics such as prior work experience, attendance at a particular high school, and GPA may all convey useful signals to both employers and job-seekers.

In this paper, we provide evidence on the role of information frictions in a large U.S. youth labor market. We partner with the New York City Summer Youth Employment Program (NYC SYEP), which employs city youth to work over the summer, to run a large-scale field experiment. The intervention provided a random subset of participants with signals about their skills that they could share with potential employers: personalized letters of recommendation from their SYEP supervisors. To test whether these signals improve employment and earnings, we follow study participants for four years in administrative unemployment insurance data from the New York State Department of Labor. To identify whether changes are driven by supply-side or demand-side responses, we invite a subset of study youth to apply to a job posting and measure application rates and confidence. And to test for impacts on educational outcomes, which could be directly affected by the letters or indirectly affected through changes in labor force involvement, we follow school-aged youth for four years in data from the NYC Department of Education.

Across a pilot after the summer of 2016 and a full-scale study after the summer of 2017, a total of 43,409 SYEP participants are in our main study sample.[2] To make letter production on this scale feasible, we invited program supervisors to complete a survey tool, developed by our research team, that automatically turned survey responses on individual participants into full-text letters of recommendation. When supervisors agreed to produce a letter of recommendation and provided high enough ratings of the youth worker, we sent that treatment youth a digital copy and five hard copies of the letter, which included, among other things, the supervisor's overall assessment of the youth worker and descriptions of relevant soft skills such as communication, reliability, and initiative.

The availability of a personalized letter of recommendation produces sizable labor market impacts. Being sent a letter increases the likelihood that a young person is employed by over 3 percentage points in the year after receiving the letter, a 4.5 percent increase relative to the 70 percent of their control group counterparts who work.[3] Employment effects fade out over the 4-year follow-up period, but not because controls completely converge: Effects on earnings grow monotonically from $150 (4.0 percent) in year 1 to $546 (5.3 percent) in year 4. Overall, sending a letter increases cumulative earnings over four years by $1,349 (4.9

2. Our empirical strategy involves stacking panels for the two cohorts, so youth can appear in the data more than once. In total, we have 43,409 observations on 41,633 unique individuals.

3. This effect is 250% as large as prior estimates of the effect of the summer program itself on employment. Gelber, Isen, and Kessler (2016) finds that the NYC SYEP increased employment by 1.2 percentage points in the post-program year, primarily by encouraging youth to participate in SYEP again.

percent). While we cannot separate hours and wages in our data, we show that treatment youth find jobs faster, work in higher-paying industries, and have longer job spells. All these results suggest that the information in the letters increased job and match quality, even conditional on working.

That simply providing a one-page letter about an individual's skills improves employment in the short term and earnings in the long term—totalling almost $12 million in additional earnings among the treatment group—suggests that information frictions significantly hamper youths' labor market trajectories. We conduct two additional sets of analyses to understand why.

First, we assess job-seeking behavior among a subset of our sample. A few months after distributing letters to treatment youth, we invited 4,000 treatment and control participants to apply for a short-term job working for us remotely. Treatment youth were 267% more likely to submit a letter of recommendation as part of their application (4.5 percent of control applicants and 16.5 percent of treatment applicants included a letter in their application), suggesting that our letter distribution translated into substantial differences in the application packages employers actually saw. But the treatment group was no more likely to apply for our job and no more likely to check a box asking to be considered for a more-selective, higher-paying job opportunity. We also find that the letters did not generate observable differences in who applied, and that changes in outside labor market involvement do not appear to explain the lack of a supply-side response. The similarity in application behavior among treatment and control youth suggests that the letters work on the demand side by changing how employers view applicants, rather than increasing motivation, job search, or confidence on the supply side.

Second, to better understand the demand-side response to the letters—in particular, to assess whether employers react to the substantive signals in the letters or just the increased salience of an application when it includes a letter—we look at treatment heterogeneity by youth ability, measured by overall supervisor rating (which we collected for both treatment and control youth). Using the control group, we show that these ratings are predictive of future earnings and educational achievement, even controlling for other covariates. This pattern suggests that the information in the letter is conveying a real signal about future productivity that might otherwise be hard for employers to observe. We then show that lasting employment and earnings increases are concentrated among youth who are more highly rated by their SYEP supervisors. Providing letters to low-rated youth generates no earnings increase and only a temporary increase in employment, driven entirely by employment with the city agency that runs the SYEP. In contrast, providing letters to high-rated youth generates higher earnings for the next four years with employers outside the city's youth development agency. That employers respond to letters about high-rated youth, but not low-rated youth, underscores that they are not just reacting to the presence of a letter, but rather the informational signal contained in the letter.

Finally, we test for treatment effects on educational outcomes. Letters could have a direct educational effect if youth also face information frictions at school such that showing the letter to teachers or guidance counselors changes the way they engage with students. Prior work has shown that teachers' and other adults' beliefs about young people directly affect their outcomes, even when the information that changed those beliefs is fictitious (Rosenthal and Jacobson 1968; Bertrand and Duflo 2017). Letters could also have an indirect effect on educational outcomes, especially if working during high school crowds out time spent in school. There is a general consensus in the literature that working a small amount has a weakly positive effect on schooling, but that working more than 20 hours is harmful (Buscha et al. 2012; Vammen Lesner et al. 2022; Staff, Schulenberg, and Bachman 2010; Monahan, Lee, and Steinberg 2011; Baum and Ruhm 2016; Ruhm 1997). However, it is difficult to isolate exogenous variation in term-time employment among high school students. And prior work has often been limited to measuring final educational attainment in surveys rather than school performance or time-to-completion. This leaves open the possibility that pushing students on the margin of dropout into the labor market could harm their educational progress—at least outside of a setting that mandates continued school enrollment as a condition of receiving a term-time job (Le Barbanchon, Ubfal, and Araya, forthcoming).

For the nearly 20,000 youth in our study who we observe in New York City public high schools, we find some indication that, by pulling people into the labor market, letters of recommendation slow down—but do not appear to stop—high school graduation. The effect is clearest among the students we would expect to be most marginal, those with below-median GPAs in the pre-randomization year. That subgroup also has the biggest change on the employment margin, accompanied by declines in enrollment and GPA during the year letters were distributed. The welfare implications of pulling these youth out of school and into the labor market depend on how long earnings benefits persist and how those benefits compare to the cost associated with a longer time spent in high school (and some students are still too young to observe final graduation outcomes, so longer-term follow-up is needed).

Most broadly, our study contributes to the literature exploring how information changes the behavior of employers and workers in the labor market. In response to fictitious applications in audit studies (Agan and Starr 2018; Kaas and Manger 2012) and policy changes in the labor market (Bartik and Nelson 2019; Doleac and Hansen 2020), employers show less discrimination when they have more information about adults. Firms' use of temp agencies or other intermediaries, screening tests, and referral bonuses suggests that employers are willing to pay to elicit particular kinds of signals for certain types of job applicants (Autor 2001; Hoffman, Kahn, and Li 2018; Pallais and Sands 2016). Job seekers also adjust their search strategies in response to tailored occupational information and to their own changing reference points over time (Belot, Kircher, and Muller 2019; DellaVigna et al. 2022).

More specifically, we expand the literature on the impacts of labor market signals about worker ability. Pallais (2014) generated seminal evidence about how uncertainty generates inefficient hiring on oDesk, a problem that gave rise to oDesk's intermediary agencies (Stanton and Thomas 2016). Pallais (2014) finds that close-to-anonymous workers with no prior experience on the platform benefit from being hired and publicly rated, while those with prior work experience benefit from more detailed reviews over the next two months. We extend the idea of public performance signals to show that they also have important impacts—over at least four years—in a state-wide labor market, and in a setting where workers can endogenously choose whether to share the relevant signal.

While we are the first paper to test the impact of reducing information frictions on labor market trajectories in a developed economy—where there are many other available signals of applicant quality—our work relates to several recent studies in developing countries' labor markets. These experiments use skill certificates and performance review templates to convincingly show that ability signals shared with employers and/or job seekers can improve labor market outcomes in low-information environments (Abebe et al. 2021; Abel, Burger, and Piraino 2020; Bassi and Nansamba 2021; Carranza et al. 2020). In addition to investigating a very different labor market setting, we complement this important work in two other ways. First, rather than focusing on those who have completed their schooling and are seeking full-time work, our study includes students, which allows us to provide new experimental evidence on substitution decisions between work and school. Second, our setting allows us to use administrative earnings data rather than self-reported survey outcomes. Although this means we can not observe hours or wages, we are able to measure all formal sector employment, separately investigate results by employer and industry, and track participants in every job over a four-year period. Measuring the full labor market trajectory is crucial to identifying if the market finds other ways to learn about the productivity of control workers (and if so, how quickly). It is also central to assessing whether employers inefficiently react to the signals (e.g., if, based on their prior experience with applicants who bring letters or skill certificates, they take the information as a more positive signal than it actually is). Inefficient belief updating could result in hiring mistakes and additional churn that would be difficult to observe in point-in-time survey data.

It is important to be clear that our conclusions are about the role of information frictions in preventing young people from obtaining successful job matches. Our findings do not necessarily imply that broader distribution of recommendation letters or other signals would be welfare enhancing. In our study, supervisors decide who to rate, negative letters are disallowed, and workers decide whether to share the signal with potential employers. It seems likely that forcing broader letter distribution would generate more negative signals, which would likely have different effects than those we estimate here (i.e., our LATE is not the ATE). Beyond treatment heterogeneity, the impact of scaling up efforts to facilitate letters of recommendation—or other credible signals—will depend on general equilibrium effects that we cannot directly measure within our study. The theory literature makes clear that the welfare effects of expanding signals in general equilibrium could increase overall employment by helping employers fill vacancies they would otherwise have left open in the face of too much uncertainty, as in Pallais (2014).[4] But it is also possible that youth with recommendation letters simply displace those without them (although this is unlikely to have happened within the context of our control group, given that there are about one million 15- to 24-year-olds in the NYC labor market and we sent fewer than 9,000 letters across two years). That said, even the welfare implications of full displacement are not obvious, since policymakers may value potential distributional changes or efficiency gains from better matches, even if there were no net change in employment.

Additional research on exactly how letters change employers' decision-making and who might be displaced by the new hires would help to predict the welfare consequences of scaling up efforts to facilitate credible productivity signals. For now, this study provides new evidence on the role of information frictions in constraining young applicants' labor market success, which could limit the impact of programs designed to improve their skills and future labor market outcomes. We establish that reducing these frictions by providing credible signals that applicants can use to communicate to employers about their strengths can significantly improve labor market trajectories.

# 2 Setting, Experiment, and Data

## 2.1 Setting

We partner with the New York City Summer Youth Employment Program (NYC SYEP), implementing our experiment with youth who participated in the summer of 2016 or the summer of 2017. The NYC SYEP is administered by the NYC Department of Youth and Community Development (DYCD). Since a post-Great Recession minimum enrollment of 29,416 youth, enrollment grew steadily to nearly 70,000 youth in 2017. In our program years, the NYC SYEP provided youth with six weeks of paid work during July and August.

4. It could also encourage youth to work harder when they know letters may be forthcoming.

All NYC residents aged 14–24 were eligible to apply for the program, though 40% of eventual participants were aged 16–17. Participants in the program were provided with jobs in the private sector (45%), at non-profits (41%), and with public sector employers (14%). The NYC SYEP directly pays youth for their work with their matched employers at the New York State minimum wage ($11.00/hour in 2017). Youth payroll totaled $83 million in 2017, or roughly $1,200 per youth participant, with a total program cost of $127 million. Over 80% of this cost was funded by the City of New York, with a majority of its remaining funding coming from New York State (see *SYEP Annual Summary* 2017).

Partnering with the NYC SYEP provides an ideal environment to assess the role of frictions in the youth labor market. SYEPs are popular and widespread social programs that provide paid work to youth—often low-income and minority youth—during the summer months, and the NYC SYEP is the largest program in the country (Heller and Kessler, forthcoming). For about half of program youth, SYEP participation is their first experience in the labor market. Consequently, SYEP participants are representative of the groups likely to face informational barriers in their attempts to capitalize on early work experience. Indeed, while SYEPs improve important outcomes including criminal justice involvement and mortality, multiple randomized controlled trials suggest they do not have consistently positive average effects on future employment (Davis and Heller 2020; Gelber, Isen, and Kessler 2016; Modestino 2019); whether information frictions constrain training programs' benefits is an open question.

## 2.2 Letter of Recommendation Experiment

We received SYEP data from DYCD on a subset of participants from the 2016 NYC SYEP (n=16,478) and all of the participants in the 2017 NYC SYEP (n=66,763). The program data identified each youth's summer work site and the supervisor or supervisors for the youth at that work site. Using these data, we limited our sample in several ways. First, since we needed to contact supervisors to ask them to complete the letter of recommendation survey, we excluded youth supervised by someone without an email address in the data. Second, we excluded some youth at large work sites to avoid making the survey unmanageable for a single supervisor. In particular, if any supervisor was linked to more than 30 treatment youth, then we randomly selected 30 treatment youth to be included in the survey. We applied the same restriction for the control youth in the survey.[5] In total, this left a sample of 69,222 SYEP participants who were included on at least one survey. Figure 1 traces

5. To ensure that neither the treatment nor control group exceeded the 30-person-per-survey limit, we randomly assigned treatment and control status prior to making these sample restrictions. Since youth were randomly selected to be excluded, random assignment is still only a function of random variables.

through this and the subsequent steps of how youth moved through the study.

To generate recommendation letters, we built a survey tool that sent a personalized survey to each supervisor asking about the youth who they supervised that summer (i.e., the youth linked to them in the DYCD data).[6] The email inviting each supervisor to participate explained the letter of recommendation program, included a link to the personalized survey tool, and encouraged them to participate (a sample of the email from 2017 is shown as Appendix Figure A.1). Supervisors were given approximately two weeks to complete the survey, and we sent up to two reminder emails to supervisors who had not yet completed it. For the 2016 cohort, we emailed 3,297 supervisors at the end of September (initial emails went out on 09/29/16). For the 2017 cohort, we emailed 11,877 supervisors in October (initial emails went out on 10/12/17).

The survey began with a brief explanation for supervisors that if they rated a youth positively enough, their responses to the survey questions might be used to construct letters of recommendation. A link to an example letter was provided to aid in the explanation. Respondents were then asked to confirm that they had been a SYEP supervisor during the preceding summer (see screens at the start of the survey in Appendix Figure A.2). Once a respondent confirmed being a supervisor, they were shown the list of treatment youth linked to them in the DYCD data, listed alphabetically by last name.[7] Supervisors selected which youth they had directly supervised and were asked a set of questions about each selected youth in a random order. The survey asked supervisors for an overall rating of the youth's performance and whether they would be willing to answer questions that would turn into a letter of recommendation for the youth (see Figure A.2 for screenshots of the survey). If they were willing, they were also invited to include their contact information on the letter of recommendation to serve as a reference (97 percent of eventual letters included contact information). They then rated the youth on several attributes, shown in Figure 2.

After the supervisors answered questions about treatment youth, they were asked one question each about control youth—the same question about the overall rating on the youth's

6. The data did not link every youth to a single supervisor. Sometimes, multiple supervisors were listed for a single work site, such that it was not clear which youth reported to which supervisor or if a youth reported to multiple supervisors; in these cases, we assumed the latter for the purposes of constructing our survey tool. Consequently, youth could be listed on more than one survey. Sometimes, a single supervisor was listed for multiple work sites. If the names of the work sites suggested they might be connected (e.g., multiple branches of the same store), we treated them as one work site for the purposes of constructing the survey tool. In the survey, we asked supervisors to confirm the youth that worked for them and to provide the names of others who might have supervised youth so we could include them in the letter of recommendation program as well. If more than one supervisor rated a young person, we generated the letter from the survey with the highest rating, breaking ties by prioritizing letters that included employer contact information, and then those with the most positive responses about the youth.

7. Note that confirming one's identity and position as an SYEP supervisor, prior to viewing treatment or control youth, is how we count "starting" the survey, a definition that is relevant below.

performance—all on one screen (see Appendix Figure A.3). They were told that these youth would not be included in the letter of recommendation program. A total of 5,854 supervisors (39 percent of all supervisors we emailed) opened the survey and confirmed that they had supervised SYEP youth during the preceding summer. In total, 43,409 young people were on a started survey, 29,887 (69 percent) of whom were given an overall rating by employers.

The software we built for this project converted the supervisors' survey responses on treatment youth into formatted letters of recommendation populated with sentences for each youth attribute. For each positively rated attribute, the letter included a dynamically constructed sentence. For example, if in response to the question "How was $<$ *youth name* $>$ at communicating?" the supervisor selected "Very effective," a sentence would appear in the letter that read: "$< Youth\ name >$ was a very effective communicator." Whereas, if the supervisor selected "Not effective" or "Somewhat effective" in response to that question, the sentence about communication would not be included in the letter.

We assigned each attribute to a potential paragraph. If the supervisor rated the youth positively enough on enough attributes to construct a particular paragraph, the paragraph was included in the letter. As long as two paragraphs could be included, the letter was generated for the youth. This procedure ensured that any letters of recommendation our survey tool generated had enough positive things to say about the youth to provide a positive letter that would not be too sparse. Our software produced letters of recommendation as PDFs on official DYCD letterhead. The letters ended with "Sincerely," followed by the name of the supervisor and work site. A short note in the footer of the letter described our letter of recommendation pilot program. Figure 3 shows a sample letter.

In total, we generated and sent 8,780 letters (1,805 in 2016 and 6,975 in 2017). We uploaded digital copies of these letters to Dropbox with a link sent to the youth for whom emails were known (1,737 in 2016 and 6,720 in 2017).[8] In addition, we mailed five physical copies of the letters via USPS to each youth along with a cover letter providing context and suggested uses for the letter (see Appendix Figure A.4 for a sample cover letter; similar text was sent to youth via email along with the link to the soft copy of the letter).[9] All letters of recommendation were sent in time for winter holiday hiring in the year after SYEP participation (letters were sent to youth in early-December 2016 for the 2016 cohort and in mid-November 2017 for the 2017 cohort).

8. About 56 percent of letter recipients clicked the link in their email to view the letter digitally. Many SYEP youth create an email solely for the purpose of the online SYEP application and then abandon it, so some letter recipients may not have seen the email containing the link to the digital copy of the letter.

9. Of the 8,780 sets of letters mailed to youth, 127 were returned as undeliverable.

## 2.3 Job Application Data

To understand the mechanisms through which letters of recommendation might impact labor market outcomes of treatment youth, we advertised a job to a subset of the youth in our data, solicited job applications, and hired youth ourselves. We composed a job listing for a short-term, flexible, and remote paid job, emailed the job listing to 4,000 randomly selected subjects from our 2017 cohort, and observed their job application behavior. The sample was evenly split among treatment and control youth from the letter of recommendation experiment who also had an email address in the data so we could send them the job application.

The job was described as being with a professor at the University of Pennsylvania who was looking for former NYC summer job participants for a short-term and flexible job. The job description highlighted several qualifications: "responsible," "self-motivated," having an "enthusiastic approach," and offered compensation of \$15/hour. A link to an application with a deadline to submit was included at the bottom of the job description (see the email invitation sent to youth with the job description in Appendix Figure A.5).

Youth who clicked the link in the email were taken to a job application that asked a few standard contact, background, and employment experience questions. We test for treatment effects on job search behavior using whether youth click the link and whether they apply for the job. Our application also provided an optional space to upload up to three "supporting documents (e.g. resume or other documents that might strengthen your application)." The application did not explicitly mention uploading letters of recommendation, but it would have been easy for youth to upload the soft copy of the letter of recommendation provided to them in our experiment (see the screenshot of this prompt in Appendix Figure A.6).[10] This upload interface allowed us to measure whether youth provided supporting materials—including a letter of recommendation—with their applications and to assess whether this differed across treatment and control youth.

Finally, to assess the confidence of youth in our study, we gave applicants the opportunity to check a box on the application to be considered for a more selective, higher-paying position (\$18/hour) that required a stronger application. The application made clear that being considered for the more selective position would not affect their chances at being selected for the regular job.

All those who submitted an application that included their name, email address, and at

10. We intentionally avoided explicitly mentioning a letter of recommendation to see if youth in our study would choose to upload a letter without a specific prompt to do so. We saw this as realistic to job applications in practice where a youth could choose to provide a potential employer with a letter of recommendation even if one was not specifically requested.

least 1 additional field were hired.[11] The job itself was an online survey of multiple-choice questions. These questions asked youth about their experiences job-seeking and considering college, as well as about their career and education goals. At the end of the survey, there were free-response questions about the youth's experience in SYEP.[12] Workers were instructed to finish everything they could within a two-hour time frame. All youth who initiated the job-task (n=227) were paid for two hours of work via a mailed, pre-loaded debit card (so our job does not appear in the administrative data on employment and earnings).

## 2.4 SYEP Administrative Data

Administrative data from the NYC SYEP comes from the NYC DYCD, which runs the program. We received data on a subset of participants of the 2016 NYC SYEP and all participants of the 2017 NYC SYEP. The data on SYEP participants include identifiers (e.g., name, date of birth, and social security number) that allow us to match to various data sources; demographics (e.g., self-identification of gender, race, and pre-SYEP education status) that allow us to test for balance and treatment effect heterogeneity; and contact information (e.g., mailing address and email address) that we used to send letters of recommendation to treatment youth. We define racial/ethnic categories based on the self-reported categories in the application, making the classifications mutually exclusive (e.g., "White" only captures non-Hispanic Whites). We also received information on the work site where the youth worked for the summer and information about the supervisors at that work site, including name and email address. We use the information on work site and supervisor to send the letter of recommendation surveys, as described above.

## 2.5 NYS Department of Labor Data

We obtained earnings and employment data from the New York State Department of Labor (NYSDOL). Data come from NYSDOL's quarterly Unemployment Insurance (UI) dataset, which covers formal sector employment, excluding self-employment or farming income. The data include employer name, employer FEIN, employer address, employer NAICS, and amount paid in each quarter. NYSDOL analysts matched SYEP participants to UI data using social security number. When multiple profiles in the NYSDOL data shared the same social security number, we used name to disambiguate the UI data. In total, 99.3 percent of SYEP youth in our letter of recommendation experiment were matched to the NYSDOL

11. To ensure our hiring for the more selective job was incentive compatible with our instructions about higher selectivity, the youth needed to click the box asking to be considered and needed to complete one or more of the free-response questions in addition to fulfilling the requirements for the standard job.

12. Youth hired for the more selective job were asked additional free-response questions that required more thoughtful consideration.

data with no difference between treatment and control youth ($\beta = 0.001, p = 0.209$).[13]

We have data from Q1 (January–March) of 2010 through Q3 (July–September) of 2021. This window provides considerable baseline data as well as four years of outcome data after letters were sent to SYEP participants in our treatment group for each study cohort.[14]

## 2.6 NYC Department of Education Data

Education data come from the NYC Department of Education (DOE).[15] The DOE used name, date of birth, and gender to perform a probabilistic match between our study sample and their records between the 2015–2016 and 2020–2021 school years, inclusive. SYEP applicants fail to match because they never appear in the DOE system (e.g., always attended private school), matched to more than one student record (DOE treats multiple matches on the same name and birth date as a non-match), or because typographical errors or name changes prevented identifying a study participant's education records. Overall, 88 percent of our sample matched to a DOE student record, with no treatment-control difference in match rates ($\beta = -0.003, p = 0.359$). Within the sample that matched to a DOE student record, 7,642 had no active enrollment within our 2015–2021 data. These students were largely old enough to have left school prior to 2015 (their average age at randomization is 19.7), although some may have transferred to private or non-NYC districts prior to the start of our data. This leaves 69.9% of our sample with at least some education information in the data, with no treatment-control difference ($\beta = -0.003, p = 0.442$).

# 3 Method of Analysis

This section discusses how we perform the analysis in this paper. In Section 3.1, we describe our sample definitions and our outcomes of interest for each data source. In Section 3.2, we describe our empirical approach, including our regression specifications. In all sections, we note cases where we deviated from our pre-analysis plan with accompanying explanations for these choices.[16]

13. In theory, everyone in our data should have matched to the data, since they were all listed as a SYEP participant during the summer prior to the program. Some of the non-workers may not have matched to the UI data despite having worked due to typographical mistakes or incorrect SSNs. Others may not have ever been paid by SYEP despite being listed as a participant in their data, and so not actually have received any wages to be reported to the UI system.

14. Letters were sent in Q4 (October–December) of 2016 or 2017, depending on cohort. Consequently, we have additional quarters of data for the youth in the 2016 cohort, but we limit the analysis to the period we can observe for full years for both cohorts.

15. At the request of the data provider, when we merge DOE data with the rest of our study data, we exclude the self-reported citizenship status that appears on the SYEP application, so that education outcomes are never linked to citizenship status. SYEP application data also provides spotty information on whether youth live in public housing or are on public assistance; those are also never linked to DOE data.

16. The pre-analysis plan can be found at https://osf.io/8zwdr/

## 3.1 Sample Definitions and Outcomes

### 3.1.1 Labor Market Sample

Our main sample to explore labor market outcomes consists of the 43,409 SYEP participants who were on a survey that a SYEP supervisor started (i.e., the SYEP participant appeared on at least one survey in which the supervisor clicked the link inviting them to take the survey and confirmed on the first page of the survey—prior to viewing which youth were on the survey or what their treatment status was—that they supervised youth that summer). This excludes the 25,813 youth who were randomized and placed on a survey that no supervisor ever opened.

We pre-specified this subsample of youth on a started survey as a key sample of interest, because neither treatment nor control youth on *unopened* surveys could have actually received treatment. This kind of non-compliance mechanically reduces statistical power and is orthogonal to treatment status, so we focus on the subsample with a first stage of 0.404 (rather than the first stage of 0.254 when we include youth on unopened surveys).[17] As a result, the treatment effect of receiving a letter of recommendation in our main sample is representative of the population of youth whose supervisors both had an up-to-date email address in the DYCD data and were willing to click on an invitation to participate in the letter of recommendation program. The estimates from this sample of youth almost certainly differ from the treatment effect on the broader sample of all SYEP youth, because different types of youth are placed into jobs with different types of supervisors, and supervisors select into responding.[18]

Since supervisor non-response was driven by an inability to reach supervisors by email or by a lack of supervisor interest or capacity to complete the survey, limiting our analysis to this sample does not interfere with the integrity of random assignment (i.e., until the supervisors reached the substantive survey questions, they had no way of knowing which youth would be included in the survey or which youth would be in the treatment or control groups).

17. While we pre-specified this subsample as a key sample of interest, our main sample included all SYEP participants that we randomized, because we did not anticipate that only 39% of supervisors would open the survey and that such a large fraction (i.e., over one-third) of the sample would be on an unopened survey. For completeness, we present and discuss results for this larger sample in Appendix Section A.6; Table A.26 shows main labor market results are quite similar, but slightly less precise. We choose to emphasize the results from our smaller sample in the main text, because the power gains from focusing on this subsample give better insight into the effect of the letter of recommendations on the sample of youth who might actually have been eligible to be treated, given the actions of their supervisors.

18. Appendix Section A.6 shows that youth who were on unopened surveys are indeed observably different than the youth in our control group of opened surveys on demographics and employment outcomes, although not in their likelihood of applying to our job posting. As such, it is plausible that forcing supervisors to rate youth would have somewhat different effects than those we estimate here.

As discussed below, Table 1 shows that our main sample is balanced across treatment and control youth.

#### 3.1.2 Labor Market Outcomes

We pre-specified a primary focus on annual earnings, winsorized to deal with outliers, and an indicator for any employment as a secondary outcome. For robustness, we also show raw earnings.[19] Results based on alternative methods of adjusting for skewness are presented in Appendix Section A.1.1. Our main analysis shows employment and earnings in each of the 4 years after randomization and counts the quarter the letters went out—the fourth quarter of the year of program participation—as the first quarter of the year (so each year is from October 1st to the following September 30th). We also show results cumulatively across all 4 years of follow-up. Note that the COVID-19 pandemic started in year 4 for the 2016 cohort and in year 3 for the 2017 cohort.

We also pre-specified exploratory analyses on: (1) the number of jobs and length of jobs to assess job stability and match quality, and (2) the industry of employment to assess whether letters help youth find jobs in which they now have experience (i.e., those over-represented in SYEP jobs) or whether the letters help market youths' skills to the higher-paying industries that are under-represented in SYEP jobs (see a discussion of these industry definitions in Gelber, Isen, and Kessler (2016)). For (1), we define a job spell as all consecutive quarters worked at the same employer. Other outcomes related to spell length and industry are discussed in Appendix Sections A.1.2 and A.1.3.

#### 3.1.3 Job Application Sample

We randomly selected 2,000 control youth and 2,000 treatment youth from our main 2017 cohort who had email addresses in the SYEP data to invite to apply to our job application.[20] This subsample is balanced on observables (joint test of treatment-control difference: p=0.219).[21]

19. To prevent too much leverage from a single outlier, the raw earnings regressions top code one observation that includes over $3 million in a single quarter to the next highest raw earnings amount in the data. The adjustment takes the yearly total for year 2 for this person from just under $3.2 million to just under $214,000.

20. We also invited 1,000 youth from unopened surveys (i.e., outside of our main sample) to ensure that job application behavior was not dramatically different for the youth excluded from our main sample.

21. Despite the overall balance, we note that the treatment group in this subsample is significantly more Hispanic by chance (33 percent in the treatment group versus 29 percent in the control group, p=0.01). As we show in Appendix Section A.3, labor market impacts for Hispanic youth are larger than for other groups. As a result, the point estimates for employment and earnings are somewhat larger for this sample.

### 3.1.4 Job Application Outcomes

For the sub-sample of individuals we randomly selected to receive our job application advertisement, we pre-specified three key outcomes: whether someone applied, whether they uploaded a letter, and whether they checked the box to apply to a more selective job as a measure of confidence. Observing whether there is a treatment-control difference in application rates helps us to test whether there is a supply-side job search response behind any potential changes in labor market outcomes. The proclivity to opt into consideration for the more selective job tests for treatment-control differences in self-efficacy and motivation or confidence in their likelihood of success on the labor market. Whether applicants uploaded a letter provides a measure of how much letter use actually changed in job applications.

We also report two additional outcomes to provide a more complete picture of job application behavior: whether someone clicked the link to view the job application (regardless of whether they applied), and whether someone uploaded any file (e.g., CV, transcript, or anything else) in support of their application.

### 3.1.5 Rated Youth Sample

To test whether the letters convey a substantive signal about worker productivity, rather than just making applications more salient, we report labor market impacts separately based on how supervisors rated an individual's overall performance. Employers were asked about the overall performance of both treatment and control youth on a seven-point scale. We split the sample into those with low overall ratings (categories 1–4: "Very Poor," " Poor," "Neutral," and "Good") and high overall ratings (categories 5–7: "Very Good," "Excellent," and "Exceptional").

Unlike our main sample, however, there is the potential for selection into who receives a rating based on supervisor behavior in the survey. Because the survey was designed to maximize the number of letters generated, treatment youth were listed first, along with a longer, multi-page set of questions on each youth; control youth were all listed at the end of the survey on a single page, with check boxes that allowed the supervisor to quickly answer the single overall quality question about each control youth. The different positioning and survey content for treatment and control youth could change the probability a supervisor rated a particular youth. Additionally, supervisors were told (and could decide whether) their responses would be turned into a letter for treatment youth, but not for control youth. The possibility of sending a letter may itself lead supervisors to make different decisions about whether to rate a youth or which rating to give. Because of both differences, we would not necessarily expect the distribution of treatment and control youth to be identical conditional on having a rating or receiving a particular rating.

In fact, treatment youth are significantly less likely to have received a rating than control youth (66 versus 71 percent, p<.01), and the distribution of ratings is somewhat different for treatment and control (see Appendix Figure A.9). There is some indication that this is driven in part by supervisors being more hesitant to give low ratings when a letter might be produced than when they knew it would not, as the distribution of baseline characteristics is nearly statistically different across treatment and control for youth receiving a low rating (p = 0.101, see Table A.5). Because we test for treatment effect heterogeneity across low-rating and high-rating groups to assess the signaling value of the letters, this kind of selection within a rating group could potentially bias our results.

To minimize the role of selection introduced by whether a youth is rated, when we report treatment effects by ratings, we focus on the sub-sample of youth who were on a survey in which the supervisor rated every treatment youth and every control youth in the survey. There are 13,911 youth who were on such a survey (4,301 with low ratings and 9,610 with high ratings). Since everyone is rated, these surveys leave no room for treatment-control differences in who is rated within the survey. In this group, treatment youth are only 0.8 percentage points less likely to appear on a completed survey overall (31.63 versus 32.46 percent, p=0.066), a small difference that might arise because it is easier to fully complete a survey with relatively fewer treatment youth. That said, the share of treatment youth on each survey is a function of random variables and, within both the low-rated and the high-rated youth of this sub-sample, observables are jointly balanced across treatment and control (see Appendix Table A.6). Appendix Section A.2 shows that even without this sample restriction, labor market results by rating are relatively similar when using all youth with a rating.

#### 3.1.6 Education Sample

Because we knew much less about what education data would be available to us at the time of pre-specification, the education analysis is where we deviate most from our pre-analysis plan.[22] As reported above, about 70 percent of our sample has any active record in the DOE data during the period we observe (2015–2021). In practice, however, many of these students either graduated or left school prior to our 2016 and 2017 study years. In addition,

22. We initially expected to use an index that included days present, an indicator for graduating or still being in school, GPA, and standardized test scores when available, plus a separate outcome measuring post-secondary enrollment. In practice, many elements of this index are missing for multiple reasons. Many students are not in school to have attendance, or they attend a school (including charters) where DOE does not share records; we do not have standardized test scores in the data (except for the selected group that takes Regents exams); and DOE measures graduation and college enrollment only for particular cohorts at particular times. Consequently, instead of forcing different patterns of missing outcomes into a single index, we instead present results separately for the outcomes we have.

while charter school students do appear in DOE data as having active records, DOE does not share with outside researchers any information about school engagement, performance, or graduation for charter school students.

We wish to avoid missing data from students who had already left school, transferred, or attended charter schools. But we cannot define our sample based on whether they have schooling records during outcome years, since treatment could affect enrollment. Instead, we define our high school sample using only baseline characteristics. We identify students who were in public, non-charter schools, attending grades 8–12 in the pre-randomization year, but who had not graduated by the August prior to the academic year the study took place. This is the group we would expect to see in high school records if they progressed through high school without transferring or dropping out. These restrictions exclude students outside of the DOE, pre-randomization dropouts and graduates, and students who temporarily stopped attending public school or had not yet joined the school district in the year before randomization. This education sample contains 19,714 students, with no treatment-control difference on being in this sample either overall ($\beta = -0.0003, p = 0.938$) or conditional on being matched to DOE data ($\beta = -0.002, p = 0.676$).[23]

#### 3.1.7 Education Outcomes

We define year 1 of educational data as the academic year (September–June) during which letters were distributed. Note that our data are at the annual level, but letters went out in November or December. As a result, only about six or seven months of year 1 captures post-treatment outcomes. To measure academic performance, we report whether a student was enrolled at all, the percent of days enrolled for which a student was present, and GPA in year 1. For enrollment, we assign 0s for those with no attendance, though they may have attended school outside our data coverage. For GPA, we use non-missing data only.[24]

As time passes, study youth will leave school for one of multiple reasons (graduation, dropout, or transfer). Since treatment could affect this behavior, later measures of educational performance could be differentially missing across treatment and control youth. To avoid this issue, we focus on longer-term educational attainment measures that can be assessed even for those not in school. For academic progress, we report the number of credits

23. We note that while the joint test of treatment-control differences on baseline observables is above traditional cutoffs ($p = 0.149$), there is some chance imbalance on race and pre-treatment GPA within the education data, discussed in more detail in Appendix Section A.4.1. One benefit of the post-double-selection LASSO that we use in our main regression specifications (as discussed below) is that it adjusts for chance imbalance in a principled way.

24. Having GPA data is balanced across treatment and control, $\beta = 0.0009, p = 0.845$. In addition, since there is treatment-control balance on whether someone is in the enrollment, attendance, and GPA data, alternative imputations of missing data would not change our results.

attempted and percentage of credits earned across the 4 years of follow-up data, including 0s for anyone who graduated, dropped out, or transferred. These measures give an overall sense of how long youth stayed in public, non-charter schools, and whether they failed a higher percentage of coursework.

Given the potential for labor market involvement to crowd out educational attainment, we are perhaps most interested in high school graduation. It is important to note that graduation data are not available for everyone. Per state standards, DOE only reports graduation in the academic years that correspond to a student's on-time (4th), 5th, or 6th year graduation cohort, even if a student returns to school after their 6th year. Graduation data are missing for students who transfer to a charter school; who move out of district; fall under another exclusion, such as having an individualized education plan (IEP); or who were not in a 4th–6th year graduating cohort between fall 2015 and summer 2021.[25]

We include 3 different measures of graduation. The first is an indicator for on-time (4-year) graduation. In our education sample, everyone is old enough to have observed at least their on-time graduation. The second measure is an indicator for whether someone ever graduated at any point in our outcome data. This captures later graduation, but some cohorts are not yet old enough to have reached their 5th- or 6th-year graduation date. Appendix Figure A.7 diagrams the available graduation data by grade and study cohort; about 6 percent of students in our education sample are too young to have 5-year graduation recorded, and 25 percent are missing 6-year graduation. These students will have 0s for "ever graduated," although they may still graduate in the future. Additionally, some students may take longer than 6 years to graduate, which (per state standards) is not captured in DOE data. To include information on whether these younger and older students are still pursuing a diploma, we create a third measure of "school persistence," which is an indicator for whether someone has either graduated or is still attending school in the 2020–21 academic year.

There are 865 youth in our education sample who do not appear in the graduation data, likely because they transferred out of the district or joined a different group excluded from state graduation counts after randomization. Since these individuals did not receive a diploma from NYC DOE, we assign them zeros for graduation. DOE discharge codes suggest there is no treatment effect on whether students transfer out of the district ($\beta = 0.003$, $p = 0.260$, with a control mean of 0.032). Since we do not observe graduation outside the

25. Note that the graduating cohort in DOE data is defined by the official 9th grade cohort to which a student belongs per state standards. We do not directly observe which graduation cohort students are in if they are not in our graduation records, so our education sample is defined based on pre-randomization grade rather than official graduating cohort. This means that students who transferred to other districts during the outcome period will remain in our data; we discuss their outcome definition below.

district, the balance on transfers helps to rule out the possibility of differential mobility biasing the graduation results.

Lastly, we have a measure of college enrollment. DOE captures post-secondary enrollment data at a single point in time, 6 months after a student reaches their on-time graduation date (i.e., only on-time graduates will have non-zero college enrollment recorded in the data). This information is based on data from the National Student Clearinghouse and from the City University of New York. Because of the timing of this measure, our post-secondary enrollment analysis makes one additional limitation relative to the education sample: it also excludes all pre-randomization 12th-graders from the "college analysis" sample, since their on-time graduation date makes their college outcome a baseline characteristic (measured just before our letters were distributed). There is treatment-control balance on the probability of being in this sample ($\beta < 0.0001, p = 0.998$).[26]

We define any post-secondary enrollment as whether someone is enrolled in a 2-year or 4-year institution 6 months after what would have been their on-time graduation date. We do not count participation in vocational or public service post-secondary activities as college enrollment. As with graduation, we assign a 0 from anyone who is part of the college analysis sample but missing from the post-secondary data.

## 3.2 Analytical Method

### 3.2.1 Main Analysis

We begin with an intent-to-treat (ITT) analysis by regressing each outcome variable on a treatment indicator and baseline covariates:

$$Y_{it} = \alpha + \beta T_i + \gamma X_{it-1} + \epsilon_{it}$$

where $Y_{it}$ is an outcome for individual $i$ at time $t$, $T_i$ is an indicator for random assignment to treatment, and $X_{it-1}$ is a vector of covariates measured at or before the time of random assignment. As pre-specified, we use a post-double-selection LASSO to select which covariates to include in each regression (Belloni, Chernozhukov, and Hansen 2014a, 2014b; Belloni et al. 2012).[27] We always include an indicator variable for study cohort, since randomization

26. We can additionally use the information on post-secondary enrollment to assess whether differential mobility is an issue for our labor market results, since we only observe UI data within New York state. For the subset of the sample with post-secondary data available, the records capture whether someone is enrolled in an out-of-state college 6 months after their on-time graduation date. The results show no evidence of differential mobility: treatment youth are no more or less likely to leave New York State for college ($\beta = -0.002$, $p = 0.692$, with a control mean of 0.065).

27. We implement this with the Stata commands pdslasso and ivlasso (Ahrens, Hansen, and Schaffer 2020). See Appendix Section A.5 for a list of the covariates we offer the LASSO, and for results without any

occurred separately by study year. Because 1,776 individuals appear more than once in the data, we cluster our standard errors on individual as identified by SSN in the SYEP data.

Not every treatment youth on a started survey was sent a letter, either because they were on a survey answered by someone who was not their direct supervisor, the supervisor did not want to provide a letter, or the supervisor provided ratings that were not positive enough for a letter to be sent. As a result of this kind of non-compliance, the ITT will understate the effect of being sent a letter. We also use random assignment as an instrument for whether a youth was sent a letter. Since we perfectly observe whether every youth was sent a researcher-generated letter, we can estimate this treatment-on-the-treated effect for everyone. We report control complier means as a baseline measure to assess proportional changes for compliers (Kling, Liebman, and Katz 2007). Below, we also provide a rough benchmark for the magnitude of actually using the letter, not just being sent a letter, leveraging data from our job application.

# 4 Results

## 4.1 Summary Statistics

Table 1 shows average pre-randomization characteristics for the treatment and control groups. No more differences are statistically different from 0 than would be expected by chance, nor are the characteristics jointly statistically different ($p = 0.201$). Study participants reflect the population that participates in NYC's SYEP. On average, they are just over 17 years old, about 43 percent male, largely identify as non-White (only 12.5 percent list being White on their application), and 75 percent report being in high school in the spring prior to the SYEP. About 45 percent of participants did not work prior to their participation in SYEP, but 97 percent work during the SYEP year, earning an average of just over $2,300 in that year, including their earnings from SYEP.

## 4.2 Labor Market Effects

Table 2 reports the main labor market effects. Panel A shows that being assigned to the treatment group increases employment rates by 1.3 percentage points (1.8 percent relative to the control mean of 70 percent) during the year following letter distribution. Actually being sent the letter increases employment in year 1 by 3.2 percentage points (4.5 percent relative to the control complier mean). The point estimates in the second year after letter distribution are about half as big and not statistically significant, and they continue to shrink over time.

covariates or with all covariates as robustness checks.

The fade out on this binary outcome is perhaps not surprising, since almost all control youth will eventually work in the formal labor market at least once (indeed, about 92 percent of controls work within the 4-year period).[28] But it is important to differentiate between two possible explanations for the fade out. The first possibility is that, as is common among active labor market interventions, the control group may completely converge with the treatment group. If this were the case, it would imply that while the signals in the letters speed up the process of information sharing, comparable information about control group workers becomes available rather quickly, such that all treatment effects are short-lived. The second possibility is that the information in the letters remains valuable over time, perhaps by helping improve the quality—or match quality—of initial jobs, which could have lasting effects (Kahn 2010; Neumark 2002; Oreopoulos, Wachter, and Heisz 2012). In this case, while the control group might catch up to the treatment group on the employment margin, the signals contained in the letters set treatment youth on a better trajectory, generating lasting earnings effects.

Panels B and C of Table 2 provide initial evidence that the letters set youth on a better labor market trajectory. They show annual raw earnings (Panel B) and winsorized earnings (Panel C). In the text, we focus on the latter, since it was our primary pre-specified outcome. Although the earnings effects for the early years are a bit noisy, they are substantively large and grow monotonically in both levels and in proportional terms over time. In year 1, the point estimate for being sent a letter is about \$150 (4 percent relative to the control complier mean, $p = 0.162$), which grows to \$546 (5.3 percent, $p = 0.085$) by year 4. Cumulatively across all 4 years, those sent a letter earned \$1,349 more than their control counterparts, a 4.9 percent increase ($p = 0.049$). This effect is not just driven by an increase on the extensive margin; conditional on working, cumulative earnings still increase (in untabulated results the IV estimate $= \$1,362, p = 0.058, N = 40,088$). The time pattern suggests that information frictions generate sustained harm to labor market outcomes.

Although we do not observe hours or wages in the UI records, we can use other aspects of the data to explore what is driving the earnings increase. Figure 4 shows effects by quarter and makes clear that the increase in work is not limited to summer jobs. Summers are quarters 3, 7, 11, and 15 in the figure, and results in those quarters do not look noticeably different from the other quarters. Table 3 reports on other measures of work intensity during this period. The first column suggests that those sent a letter work an additional 0.15 quarters over 4 years (a 2 percent increase), although the result is not statistically significant ($p = 0.129$). The second column shows that there is no increase in the number of

28. It is also possible some of the fade out has to do with the difficulty of finding work at the beginning of the COVID-19 pandemic, which is reflected in the downward shift of control means in year 3 at a time when more youth should have been joining the labor force as they age.

job spells. The point estimate on the number of jobs (including 0s) is positive but far from statistically significant, partly reflecting the change at the extensive margin of working at all. Conditional on working at all (the third column), the point estimate shrinks by about half to a 0.7 percent increase. The fifth column shows that, conditional on working, treatment youth find jobs sooner than controls (0.25 quarters sooner for letter recipients). Together, these results suggest that better signals help youth shorten the job search process but do not simply substitute early work for later work or generate more churn.

The fourth column documents an increase in average spell length, which suggests that the additional work is driven by successful job matches. We measure spell length by averaging across the first 3 (non-missing) job spells after randomization. We limit our attention to 3 spells in part because the average number of spells is just over 3, and in part because after 3 spells, censoring from the end of our data becomes a larger issue.[29] Average spell length significantly increases by about 0.12 quarters (3.6 percent). The longer job spells suggest that the recommendation letters increased worker and/or employer satisfaction with the job match. This finding argues against the hypothesis that employers are inefficiently updating off of the signals in the letter and instead suggests the signal in the letter is generating helpful sorting.[30]

Appendix Section A.1.3 provides some additional evidence on job type, suggesting that the signals in the letters help youth secure better jobs. Table A.3 shows that treatment youth do not just return to the agency that runs the SYEP; employment and earnings increases are concentrated among non-DYCD employers. Table A.4 uses the industry groupings from Gelber, Isen, and Kessler (2016) to show that letters seem to shift young workers into the higher-paying industries that are typically under-represented among the jobs that the NYC SYEP offers. Overall, it appears that the signals generate a long-term increase in earnings by helping young workers find better jobs sooner and stay in these jobs longer.

## 4.3 Mechanisms

### 4.3.1 Assessing Changes in Labor Supply

A key question about the observed increase in labor market success among treatment youth is whether the letters increase labor supply by increasing youth job search intensity or confidence, or whether the letters increase labor demand by changing beliefs about applicants

29. Across all 3 spells, there is no differential censoring across treatment groups; the treatment effect on the number of censored spells is -0.004 ($p = 0.340$). Appendix Section A.1.2 shows additional details about spell length and censoring among the first 3 spells.

30. If employers' prior experience with recommendation letters led them to believe that only very high-productivity workers have such signals, they might have been induced into hiring mistakes that could have generated additional churn. In practice, the longer spells suggest the reverse.

with letters or increasing the salience of those applicants among employers. By distributing our own job posting to 4,000 treatment and control youth, we are able to generate some evidence on why the letters increase employment and earnings.

Table 4 suggests that supply-side responses—increased job search, motivation, or confidence—are unlikely to be driving the labor market improvements. We find no evidence that treatment youth are more likely to click on the application link or actually apply to our posting.[31] The second column shows that 8.8 percent of the control group and 8.2 percent of the treatment group applied to our job, a difference that is not statistically significant. We also find no evidence that the letter increased confidence among applicants conditional on applying; treatment youth are no more likely to volunteer for the more selective job than control youth (see the third column of Table 4, which, adjusting for application rates, translates into 60 percent of control applicants and 51 percent of treatment applicants checking the box to apply for the more selective job).

Of course, it is possible that even though the letters did not change the rate at which young people applied to our job, they could have changed the composition of who applied. This might be the case if letters help young job seekers better target their job applications to appropriate opportunities, or if the treatment group's increased formal labor market involvement reduced their interest in our short-term, online job—despite our framing the job as flexible enough to be compatible with other work.

To assess this possibility, we test for compositional differences between treatment and control applicants on observables, and we find no clear evidence that observables are jointly correlated with treatment.[32] We also test for differences in job application behavior only among those not employed elsewhere during the quarter the job application was distributed (despite being a selected group). Even among this group, there is still no statistically significant difference in application rates or in our confidence measure ($\beta = -0.01$, $p = 0.354$ for applying and $\beta = -0.01$, $p = 0.133$ for checking the selective box).[33] We conclude that the lack of an increase in supply-side job-seeking behavior does not appear to be due to treatment changing the composition of applicants or increasing other employment. Overall, the

31. The "applied" variable here measures whether a youth entered enough information in the application for us to know who filled out the application. We define "applied" this way because we hired people even if they did not answer all the questions on the application. To actually be hired, the youth additionally needed to click submit on the final page of the application. There is also no treatment-control difference on whether youth were hired per this definition.

32. We test for differences between the treatment and control individuals who applied for our job by interacting each baseline covariate with an indicator for whether the individual applied, regressing treatment on all covariates and these interactions, and then testing the hypothesis that all interaction coefficients are jointly 0. The p-value of this test is 0.15.

33. The same is true conditional on being employed in that quarter: $\beta = 0.0008$, $p = 0.959$ for applying and $\beta = -0.01$, $p = 0.481$ for confidence.

evidence from our job application suggests that labor market improvements are coming from employers responding to letters of recommendation, not from changes in youth application behavior or confidence.[34]

As an important check on whether treatment youth actually use the letters we send them—a necessary condition for employers to be able to respond to the letters—the final two columns of Table 4 show treatment effects on the files job applicants uploaded in their application to our job posting. There is no detectable change by treatment in the probability that youth upload some form of supporting material. But there is a dramatic change in whether youth upload a letter of recommendation. Only 0.4 percent of the control group submits a letter, including zeros for those who do not apply (conditional on applying, this translates to 4.5 percent of control applicants submitting a non-intervention letter with their application). Treatment youth are two and a half times more likely to submit a letter of recommendation than the control group: 1.4 percent of all those invited to apply submit a letter (16.5 percent conditional on applying). Since about 40 percent of treatment youth actually received a letter, this implies that about 41 percent of letter recipients use them when they apply to a job (16.5 percent relative to 40 percent).

Given the lack of a supply-side response, it is possible that letters *only* work when employers see them. If so, we could use the observed rates at which letters are used in our job application as an implied first stage of letter use, providing a back-of-the-envelope extrapolation of how big employment responses would be for youth who actually use their letters. If we make the quite strong assumptions that the difference in letter use we observe in our job application applies to the entire sample, that treatment and control youth apply to jobs at the same rate, and that everyone applies to at least one job, then the implied first stage for letter use is a 12 percentage point increase (4.5 versus 16.5 percent among applicants). Scaling our main ITT effects by this first stage would in turn imply that the employment increase for those who use the letter is about 11 percent relative to baseline in the first year, with an additional $4,500 in earnings over 4 years. Of course, many of the assumptions involved in this benchmark could fail, including the exclusion restriction. The calculation is just intended to give a rough sense of how big employer responses would be for compliers in this simple case.

34. This is one key difference between our results and those in developing countries, which typically find supply-side responses to employer feedback or skill certifications (Abel, Burger, and Piraino 2020; Bassi and Nansamba 2021; Carranza et al. 2020). It may be that the higher level of unemployment in African countries (Bandiera et al. 2021) generates more discouragement that performance information can reverse, as in the Gonzalez and Shi (2010) model.

### 4.3.2 Assessing Changes in Labor Demand

The evidence so far suggests that employers are the ones responding to the signals in the letters that treated youth include in their job applications. The way that employers respond to these productivity signals is consistent with a range of models that show how employer uncertainty about applicant ability can generate inefficient hiring. In these models, young or novice applicants may remain unemployed, badly matched, or paid less than their marginal product, because only those applicants whose expected productivity exceeds an employer's cost threshold are hired or efficiently paid, where employer costs can come from risk aversion, screening costs, or the transaction costs of bad hires (Altonji and Pierret 2001; Farber and Gibbons 1996; Kahn and Lange 2014; Pallais 2014).[35] Typically in these models, unemployed but high-ability workers at the margin of being hired would like to invest in signals of their ability, which would improve market efficiency. But such workers cannot generate credible signals on their own. The existing market under-provides these signals because the employers who bear the cost of producing them do not fully internalize the gains to other employers and workers (Becker 1964), or because, in practice, novice workers under-estimate the benefits of asking for them (as in Abel, Burger, and Piraino (2020)).

Given the lack of a supply-side response, the lasting earnings effects, and the longer job spells, it seems plausible that—consistent with these models—employers are using the letters to identify high-productivity workers who they might not have hired absent the signals in the letters (e.g., because of too much uncertainty about the workers' abilities). Because our letters vary in how strong a positive signal they communicate, they can also facilitate a more direct test for whether employers' inability to identify qualified applicants—including otherwise hard-to-observe characteristics like enthusiasm and reliability—actually hinders young people's labor market trajectories.

To test the signaling mechanism, we proxy for variation in worker ability with supervisors' overall quality ratings of each youth, which we observe for both treatment and control groups. Employers can observe this variation in signal strength for the treatment group, since higher ratings correspond with a stronger introductory sentence in the letter, as well as longer letters mentioning more positive attributes. If employers are using the information in letters to update their beliefs about each individual, we would expect higher-rated youth to benefit more than lower-rated youth from receiving a letter, since the letters of higher-rated youth allow them to signal their higher ability.[36]

35. We focus on employer uncertainty here, because we do not find a supply-side response. But models incorporating uncertainty among job-seekers produce similarly inefficient hiring and match productivity (e.g., Gonzalez and Shi 2010; Mortensen and Pissarides 1999).

36. Appendix Section A.3 discusses other types of heterogeneity and their interpretations.

As a sanity check on whether ratings and letter quality could plausibly provide an accurate signal about unobserved future worker ability, Table 5 shows the correlation between supervisor ratings and future earnings within the control group, as well as the correlation of supervisor ratings with future GPA and school persistence in the education sample. Ratings are significantly correlated with all three outcomes unconditionally. And while the magnitude of the relationship gets smaller when controlling for all the other covariates we use in our regressions (including demographics, employment history, and, for the educational outcomes, prior school performance), ratings still significantly predict all these future outcomes conditional on observables. In terms of magnitude, all else equal, a one standard deviation increase in supervisor rating (1.5 additional points on the 1–7 scale) corresponds to a substantial shift in earnings and a moderate shift in education outcomes: $1,458 more in cumulative earnings, 0.45 additional GPA points, and a 0.75 percentage point increase in the probability of having graduated or continuing to attend school by the end of the data. So it appears that supervisor ratings communicate real information about expected worker productivity in their letters that might otherwise be difficult for potential employers to observe.

Table 6 shows that the impact of treatment on future labor market outcomes is larger for youth with more positive signals. The table separately estimates employment and earnings effects for youth who received low ratings (categories 1–4, corresponding to "Very Poor," " Poor," "Neutral," and "Good") and high ratings (categories 5–7, corresponding to "Very Good," "Excellent," and "Exceptional").[37] Highly rated youth were much more likely to receive a letter (81 percent versus 33 percent). So the ITT differences between the groups reflect both differences in letter receipt and differences in outcomes conditional on being sent a letter, although the substantive pattern of results is quite similar for both the ITT and TOT.

We find that the low-rated group has net employment effects close to 0 and cumulative earnings point estimates that are negative but with large standard errors. They do have a marginally statistically significant increase in employment in year 1. While this might be a chance finding given the number of hypothesis tests in the table, deeper exploration shows that this is driven entirely by increased employment at DYCD, the agency that runs the SYEP and other year-round workforce development programs (the year 1 employment impact at DYCD is a marginally statistically significant 0.027, which is comparable to the estimate reported here, which averages across both DYCD and non-DYCD employers, of

37. Note that if youth received an overall rating less than "Good," the paragraph that included text about the overall rating was not printed in the letter. Such letters could still be produced, however, as long as enough other attributes were rated positively.

0.0247). Given prior evidence that SYEP participation itself may lower future earnings by encouraging youth to work in the lower-paying industries that are over-represented in the summer program (Gelber, Isen, and Kessler 2016), the additional connection to DYCD jobs may help to explain the directionally negative earnings estimates. Overall, it appears letters that provided only a weakly positive signal did not change employer beliefs enough to help young people get or keep jobs outside of the government agency designed to support them.

In contrast, the high-rated group has lasting positive and significant employment effects, including in years 2 and 3, as well as consistently positive earnings effects that grow over time. Although the test of the difference between groups varies in its level of statistical significance across outcomes and time periods, enough of the treatment effects are statistically different between the low-rated and high-rated groups to have some confidence in the result that letters generate larger and more persistent labor market improvements for the high-rated than the low-rated youth.

One might wonder whether effects are smaller among the low-rated group simply because they choose not to use letters in their job applications. Results from our job application suggest otherwise (see Appendix Section A.2). For every 100 letters sent to high-rated treatment youth, we received 3 job applications that included letters. For every 100 letters sent to low-rated treatment youth, we received 4 applications including letters. This pattern suggests that low-rated letter recipients are not less likely to use letters when applying for jobs.[38]

It appears, then, that employers are receiving letters from both high and low rated youth and are using the substance of the letters as a signal about who is likely to be a productive employee (i.e., not just taking more notice of all applications that include letters). This result suggests that information frictions may be holding back relatively high-performing youth workers (relative to other SYEP participants) and that simple credible signals can help improve the labor market prospects of these youth. Meanwhile, the group of young people who did not impress their SYEP supervisors as much may need more intensive investments—such as improvements in their human capital—to improve their labor market outcomes.

A natural question, given that our entire sample is composed of SYEP participants, is about the external validity of our results. Are we documenting a general phenomenon about information in the youth labor market in New York City, or are we documenting that letters help overcome a particular stigma associated with SYEP participation? The answer to this question rests in part on whether the employers in our data know that youth applicants are SYEP participants, which is necessary for the stigma story. While we do not observe

38. While high-rated letter recipients apply at somewhat higher rates and use letters somewhat more often, many more of high-rated youth are sent letters than their low-rated counterparts.

that directly, we can take a hint from the applications that youth submitted in response to our job advertisement. In those applications, only 22 percent of applicants self-identify as a SYEP participant in either their list of work experience or their résumé. Given that almost 80 percent of job applicants may not appear as prior SYEP participants to employers, it seems plausible that the frictions we document are not specific to SYEP-related beliefs among employers.[39]

## 4.4 Education Outcomes

Unlike the existing experimental literature on labor market signaling in developing countries, which focuses on those seeking full-time employment, SYEPs serve many young people who are still in high school. The inclusion of school-aged youth in our sample allows us to explore two ways in which the letters could affect educational outcomes. First, given evidence that teachers' beliefs about students shape educational performance, even when beliefs are based on fictional information (Papageorge, Gershenson, and Kang 2020; Rosenthal and Jacobson 1968), letters could improve educational outcomes if teachers update their beliefs about a student after seeing the signal in the letter. The instructions we sent young people mentioned showing letters to teachers and guidance counselors for this reason.[40]

Second, pulling students into the labor market could pull them out of school. This kind of substitution has long been a concern in the literature on working during school. The consensus from natural variation in term-time work suggests that working fewer than 20 hours a week during school has weakly positive effects on education (Buscha et al. 2012; Staff, Schulenberg, and Bachman 2010; Lesner et al. 2022; Monahan, Lee, and Steinberg 2011; Baum and Ruhm 2016; Ruhm 1997). This consensus has led to a policy push to offer year-round employment opportunities to students. But analyzing students who choose to work could be different than pushing students at the margin (e.g., those who would not find work without the letter) into the labor market. To our knowledge, the only causal evidence on this question comes from a program in Uruguay that conditioned participation in a work program on staying in school (Le Barbanchon, Ubfal, and Araya, forthcoming), so it cannot speak to what youth would choose in the absence of a structured incentive to maintain school attendance.

Table 7 shows treatment effects for the 45 percent of our sample we expect to observe in high school outcome data (see Section 3.1.6 for the detailed sample definition). Overall, point

39. It is also worth noting that since the recommendation letters we distribute come on letterhead from DYCD—the agency that runs the SYEP—any stigma against SYEP or DYCD among employers would push against us finding a positive impact of the letters on employment outcomes.

40. A related mechanism would be if the signal in the letter changed the youth's beliefs about their own ability to succeed in school (e.g., if it boosted their academic confidence).

estimates suggest some declines in performance and attainment associated with the letters of recommendation, but they are not statistically significant. The one marginally significant result is a decline in on-time graduation (a 1.6 percentage point, or 2 percent, decline for letter recipients). The effects on ever graduating and school persistence (graduating or still attending) are considerably smaller, hinting that letters might slow down, but not stop, high school completion.

On one hand, we might not want to make too much of this result given that there is only one marginally significant change across 9 different outcomes. On the other hand, graduating from high school is substantively important enough that even a suggestion that youth substitute work for school is worth additional attention. Appendix Section A.4 provides evidence that this substitution is real. There we use outcomes measuring work and school jointly to confirm that it is the same group of people driving the increase in work and decline in school progress (both on-time graduation and overall persistence). We also use GPA in the pre-randomization year to focus attention on students who are most likely to be at the margin of graduating. For below-median GPA students, the letters are associated with a significant decline in year 1 enrollment and GPA, as well as a substantively large and significant decline in on-time graduation (a decrease of 5.7 percentage points, or 8 percent). These students also show the largest employment change on the extensive margin.

We should use some caution in interpreting these results, since they involve an exploratory subgroup split among the already-reduced education sample. Nonetheless, all indications are that students who are close to the margin of failing to graduate take longer to complete high school, driven by those who shift from not working to working when they receive the letter of recommendation. Longer follow-up that includes 5th- and 6th-year graduation outcomes for the youngest cohorts is needed to know for sure if this group will eventually return to school.

# 5 Discussion

Sending youth a few copies of a recommendation letter and an email with a link to that letter improves their labor market trajectories. Short-term employment rates increase by 4.5 percent, while earnings effects continue to grow over 4 years, generating a cumulative 4.9 percent increase. These are big changes for a small intervention, providing new evidence that there are information frictions in the U.S. youth labor market that credible signals can overcome. The lack of differences in job-seeking behavior among treatment youth, other than using the letter itself, suggests that employers are the ones responding to the additional information contained in the letters. And given that higher performing youth get a larger labor market benefit from the letters, it seems that employers are successfully updating their beliefs about which applicants are likely to be productive workers.

We also find that recommendation letters may lead to a decline in on-time graduation, driven by substitution toward work among lower-achieving students. To assess the welfare implications of this substitution, we would need to make some strong assumptions about how long the increase in earnings will last and how that compares to the costs of additional time spent in high school. It seems likely that the net effect may not be beneficial, especially if, once the whole sample is old enough to observe final graduation outcomes, any subgroup has left school entirely. Reducing employment frictions is most likely to have a net benefit for those who have already finished their high school careers, or at least are not currently on the margin of failing to graduate. Current policy efforts to provide subsidized, year-round employment opportunities for high school students might therefore benefit from careful targeting.

For those not on the margin of on-time graduation, it is clear that providing credible signals about existing skills benefited the individual youth in our study. This is an important insight for social programs looking to help young people capitalize on their training or skills, and potentially for other disadvantaged groups facing employer uncertainty in the labor market (e.g., the formerly incarcerated). Finding additional ways to provide personalized information about an individual worker's strengths could help improve labor market outcomes among low-income, largely minority individuals like those in our study.

That said, it is crucial to emphasize that improvements among individuals are not the same as increases in net welfare. While the small size of our control group relative to the whole labor market makes within-study spillovers implausible, we do not have the data to assess the broader societal impact of providing some youth with letters. It is not clear whether the reduction in uncertainty decreased vacancies that would otherwise have been left open, as they appear to in an online marketplace (Pallais 2014), or whether other workers were displaced (and if so, whether the resulting hires were more or less productive than the potentially displaced workers).

The possibility of displacement, along with the local nature of our effects, means that it is not clear whether scaling up the provision of credible signals would generate welfare gains. Our effects are specific to the youth who receive letters when survey responses are voluntary and responses are positive enough. Effects may differ if a broader population of supervisors or young people were pushed into participating.[41]

41. It is difficult to say from the observable differences in youth across the opened and unopened surveys whether effects would be bigger or smaller if supervisors were forced to fill out the surveys. The unopened surveys contained more White youth, for whom we observe smaller labor market effects. But they also had more youth already out of high school, which could diminish graduation crowd-out, and more youth with work experience prior to SYEP, who have directionally larger point estimates on employment and earnings,

Of course, there are several conditions under which a scaled-up version could be beneficial, even with no net employment increase and full displacement. It is possible that widespread skill signals could generate efficiency gains by helping employers and employees find better matches. Alternatively, policymakers with preferences for equity might value transferring job opportunities to those farther down the income distribution or to historically marginalized groups. Finally, there could also be general equilibrium effects on the supply side; if young people understand that they may receive helpful recommendation letters, they may work harder in their jobs, generating additional productivity as well as better letters to which future employers will respond more positively.

These general equilibrium questions are an important avenue for future work. Further research into the precise way employers update their beliefs, substitute across workers, or change the number of employees in response to reduced uncertainty might be particularly productive next steps in assessing the most effective way to leverage our findings into welfare-enhancing labor market policies. But the key conclusion from our experiment—that information frictions reduce long-term earnings trajectories, even in high-information settings like the U.S.—does not rest on anticipating the effects of signals at scale in equilibrium. Documenting the existence of these frictions is enough to confirm the role of employer uncertainty in the labor market, as well as to raise key questions about the way in which this uncertainty might be limiting the success of active labor market programs and other efforts to improve the labor market prospects of young people.

see Appendix Section A.3.

# References


Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. 2021. "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City." *The Review of Economic Studies* 88 (3): 1279–1310.

Abel, Martin, Rulof Burger, and Patrizio Piraino. 2020. "The Value of Reference Letters: Experimental Evidence from South Africa." *American Economic Journal: Applied Economics* 12 (3): 40–71.

Agan, Amanda, and Sonja Starr. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *The Quarterly Journal of Economics* 133 (1): 191–235.

Ahrens, Achim, Christian Hansen, and Mark Schaffer. 2020. *pdslasso and ivlasso: Progams for post-selection and post-regularization OLS or IV estimation and inference.* http://ideas.repec.org/c/boc/bocode/s458459.html.

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer learning and statistical discrimination." *The Quarterly Journal of Economics* 116 (1): 313–350.

Autor, David. 2001. "Why Do Temporary Help Firms Provide Free General Skills Training?" *The Quarterly Journal of Economics* 116 (4): 1409–1448.

Bandiera, Oriana, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2021. *The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda.* SSRN Scholarly Paper ID 3910330. Rochester, NY: Social Science Research Network.

Bartik, Alexander, and Scott Nelson. 2019. "Deleting a Signal: Evidence from Pre-Employment Credit Checks." *University of Chicago, Becker Friedman Institute for Economics Working Paper,* nos. 2019-137.

Bassi, Vittorio, and Aisha Nansamba. 2021. "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda." *The Economic Journal* 132 (642): 471–511. ISSN: 0013-0133.

Baum, Charles L., and Christopher J. Ruhm. 2016. "The Changing Benefits of Early Work Experience." *Southern Economic Journal* 83 (2): 343–363.

Becker, Gary S. 1964. *Human capital: A theoretical and empirical analysis, with special reference to education.* University of Chicago Press.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen. 2012. "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain." *Econometrica* 80 (6): 2369–2429. ISSN: 1468-0262.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014a. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives* 28 (2): 29–50. ISSN: 0895-3309.

———. 2014b. "Inference on Treatment Effects after Selection among High-Dimensional Controls†." *The Review of Economic Studies* 81 (2): 608–650. ISSN: 0034-6527.

Belot, Michele, Philipp Kircher, and Paul Muller. 2019. "Providing advice to jobseekers at low cost: An experimental study on online advice." *The Review of Economic Studies* 86 (4): 1411–1447.

Bertrand, Marianne, and Esther Duflo. 2017. "Field experiments on discrimination." *Handbook of Economic Field Experiments* 1:309–393.

Buscha, Franz, Arnaud Maurel, Lionel Page, and Stefan Speckesser. 2012. "The Effect of Employment while in High School on Educational Attainment: A Conditional Difference-in-Differences Approach." *Oxford Bulletin of Economics and Statistics* 74 (3): 380–396.

Card, David, Jochen Kluve, and Andrea Weber. 2018. "What works? A meta analysis of recent active labor market program evaluations." *Journal of the European Economic Association* 16 (3): 894–931.

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin. 2020. *Job Search and Hiring with Two-sided Limited Information about Workseekers' Skills.* W.E. Upjohn Institute.

Crépon, Bruno, and Gerard J Van Den Berg. 2016. "Active labor market policies." *Annual Review of Economics* 8:521–546.

Davis, Jonathan M.V., and Sara B. Heller. 2020. "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs." *The Review of Economics and Statistics* 102 (4): 664–677.

DellaVigna, Stefano, Jörg Heining, Johannes F Schmieder, and Simon Trenkle. 2022. "Evidence on job search models from a survey of unemployed workers in germany." *The Quarterly Journal of Economics* 137 (2): 1181–1232.

Doleac, Jennifer L., and Benjamin Hansen. 2020. "The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden." *Journal of Labor Economics* 38 (2): 321–374.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman. 2020. *Labor market dynamics and development.* Technical report. JSTOR.

Farber, Henry S, and Robert Gibbons. 1996. "Learning and wage dynamics." *The Quarterly Journal of Economics* 111 (4): 1007–1047.

Freeman, Richard Barry, and Harry J Holzer. 1986. *The black youth employment crisis.* University of Chicago Press.

Freeman, Richard Barry, and David A Wise. 1982. *The youth labor market problem: Its nature, causes, and consequences.* University of Chicago Press.

Gelber, Alexander, Adam Isen, and Judd B. Kessler. 2016. " The Effects of Youth Employment: Evidence from New York City Lotteries." *The Quarterly Journal of Economics* 131 (1): 423–460.

Gonzalez, Francisco M, and Shouyong Shi. 2010. "An equilibrium theory of learning, search, and wages." *Econometrica* 78 (2): 509–537.

Heinrich, Carolyn J, and Harry J Holzer. 2011. "Improving education and employment for disadvantaged young men: Proven and promising strategies." *The Annals of the American Academy of Political and Social Science* 635 (1): 163–191.

Heller, Sara B, and Judd B Kessler. Forthcoming. "How to allocate slots: The market design of Summer Youth Employment Programs." *Fair by Design: Economic Design Approaches to Inequality, Eds. S.D. Kominers and A. Teytelboym.*

Hoffman, Mitchell, Lisa B. Kahn, and Danielle Li. 2018. "Discretion in Hiring." *The Quarterly Journal of Economics* 133 (2): 765–800.

Holzer, Harry J. 1988. "Search method use by unemployed youth." *Journal of Labor Economics* 6 (1): 1–20.

Hoynes, Hilary, Douglas L Miller, and Jessamyn Schaller. 2012. "Who suffers during recessions?" *Journal of Economic Perspectives* 26 (3): 27–48.

Jovanovic, Boyan. 1979. "Job matching and the theory of turnover." *Journal of Political Economy* 87 (5, Part 1): 972–990.

Kaas, Leo, and Christian Manger. 2012. "Ethnic discrimination in Germany's labour market: A field experiment." *German Economic Review* 13 (1): 1–20.

Kahn, Lisa B, and Fabian Lange. 2014. "Employer learning, productivity, and the earnings distribution: Evidence from performance measures." *The Review of Economic Studies* 81 (4): 1575–1613.

Kahn, Lisa B. 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics* 17 (2): 303–316.

Katz, Lawrence F, Jonathan Roth, Richard Hendra, and Kelsey Schaberg. 2020. *Why Do Sectoral Employment Programs Work? Lessons from WorkAdvance.* Technical report. National Bureau of Economic Research.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Le Barbanchon, Thomas, Diego Ubfal, and Federico Araya. Forthcoming. "The Effects of Working while in School: Evidence from Uruguayan Lotteries." *American Economic Journal: Applied Economics.*

Lesner, Rune Vammen, Anna Piil Damm, Preben Bertelsen, and Mads Uffe Pedersen. 2022. "The Effect of School-Year Employment on Cognitive Skills, Risky Behavior, and Educational Achievement." *Economics of Education Review* 88:102241.

Modestino, Alicia Sasser. 2019. "How Do Summer Youth Employment Programs Improve Criminal Justice Outcomes, and for Whom?" *Journal of Policy Analysis and Management* 38 (3): 600–628.

Monahan, Kathryn C., Joanna M. Lee, and Laurence Steinberg. 2011. "Revisiting the Impact of Part-Time Work on Adolescent Adjustment: Distinguishing Between Selection and Socialization Using Propensity Score Matching." *Child Development* 82 (1): 96–112.

Mortensen, Dale T, and Christopher A Pissarides. 1999. "New developments in models of search in the labor market." *Handbook of Labor Economics* 3:2567–2627.

Neumark, David. 2002. "Youth Labor Markets in the United States: Shopping Around vs. Staying Put." *Review of Economics and Statistics* 84 (3): 462–482.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics* 4 (1): 1–29.

Pallais, Amanda. 2014. "Inefficient Hiring in Entry-Level Labor Markets." *American Economic Review* 104 (11): 3565–3599.

Pallais, Amanda, and Emily Glassberg Sands. 2016. "Why the Referential Treatment? Evidence from Field Experiments on Referrals." *Journal of Political Economy* 124 (6): 1793–1828.

Papageorge, Nicholas W., Seth Gershenson, and Kyung Min Kang. 2020. "Teacher Expectations Matter." *The Review of Economics and Statistics* 102 (2): 234–251.

Rosenthal, Robert, and Lenore Jacobson. 1968. "Pygmalion in the classroom." *The Urban Review* 3 (1): 16–20.

Ruhm, Christopher J. 1997. "Is High School Employment Consumption or Investment?" *Journal of Labor Economics* 15 (4): 735–776.

Staff, Jeremy, John E. Schulenberg, and Jerald G. Bachman. 2010. "Adolescent work intensity, school performance, and academic engagement." *Sociology of Education* 83 (3): 183–200.

Stanton, Christoper T., and Catherine Thomas. 2016. "Landing the First Job: The Value of Intermediaries in Online Hiring." *The Review of Economic Studies* 83 (2): 810–854.

*SYEP Annual Summary.* 2017. Technical report. New York City Department of Youth and Community Development.

Vammen Lesner, Rune, Anna Piil Damm, Preben Bertelsen, and Mads Uffe Pedersen. 2022. "The Effect of School-Year Employment on Cognitive Skills, Risky Behavior, and Educational Achievement." *Economics of Education Review* 88. ISSN: 0272-7757.

Van der Berg, Servaas. 2007. "Apartheid's enduring legacy: Inequalities in education." *Journal of African economies* 16 (5): 849–880.

# Tables and Figures

Table 1: Descriptive Statistics

| | Control Mean | Treatment Mean | Test of Difference |
|---|---|---|---|
| N | 21,695 | 21,714 | |
| Age | 17.2 | 17.2 | 0.641 |
| Male | 0.427 | 0.427 | 0.991 |
| Black | 0.409 | 0.411 | 0.805 |
| Hispanic | 0.289 | 0.289 | 0.944 |
| Asian | 0.129 | 0.130 | 0.734 |
| White | 0.124 | 0.125 | 0.756 |
| Other Race | 0.049 | 0.045 | 0.080 |
| In High School | 0.755 | 0.751 | 0.339 |
| HS Graduate | 0.044 | 0.042 | 0.202 |
| In College | 0.173 | 0.180 | 0.081 |
| Not in UI Data | 0.006 | 0.007 | 0.209 |
| Never Employed Pre-SYEP | 0.450 | 0.457 | 0.125 |
| Ever Worked, Year -4 | 0.153 | 0.149 | 0.225 |
| Earnings, Year -4 | 303 | 311 | 0.576 |
| Ever Worked, Year -3 | 0.266 | 0.266 | 0.866 |
| Earnings, Year -3 | 574 | 575 | 0.931 |
| Ever Worked, Year -2 | 0.437 | 0.435 | 0.648 |
| Earnings, Year -2 | 1052 | 1031 | 0.370 |
| Ever Worked, Year -1 | 0.965 | 0.966 | 0.699 |
| Earnings, Year -1 | 2334 | 2325 | 0.757 |
| No Education Match | 0.126 | 0.123 | 0.359 |
| In HS Sample | 0.454 | 0.454 | 0.938 |
| Joint F-Test | $F(24, 41632) = 1.23$, $p=.201$ | | |

Notes: N = 43,409. 390 youth missing race/ethnicity and 1 missing self-reported education status. Test of Difference reports the p-value from a regression of each characteristic on a treatment indicator, controlling for a cohort indicator and using standard errors clustered on individual.

Table 2: Labor Market Effects

| Year | 1 | 2 | 3 | 4 | Cumulative |
|---|---|---|---|---|---|
| | Panel A: Employment | | | | |
| ITT | 0.0128*** | 0.0059 | 0.0029 | 0.0011 | 0.003 |
| | (0.0041) | (0.0041) | (0.0043) | (0.0044) | (0.0025) |
| CM | 0.701 | 0.720 | 0.650 | 0.682 | 0.922 |
| Sent Letter (IV) | 0.0316*** | 0.0145 | 0.0077 | 0.0028 | 0.0071 |
| | (0.0102) | (0.0102) | (0.0107) | (0.0108) | (0.0062) |
| CCM | 0.697 | 0.728 | 0.662 | 0.697 | 0.924 |
| | Panel B: Earnings | | | | |
| ITT | 57.13 | 106.12 | 156.23 | 243.92* | 531.42* |
| | (49.29) | (77.61) | (101.99) | (136.59) | (297.78) |
| CM | 3594 | 6005 | 7457 | 10057 | 27141 |
| Sent Letter (IV) | 143.47 | 273.27 | 399.03 | 619.14* | 1357.25* |
| | (121.71) | (191.76) | (252.05) | (337.61) | (735.64) |
| CCM | 3764 | 6215 | 7579 | 10309 | 27943 |
| | Panel C: Earnings, Winsorized at 99th Percentile | | | | |
| ITT | 57.96 | 104.37 | 128.83 | 214.72* | 544.52** |
| | (43.16) | (71.94) | (96.65) | (128.38) | (277.26) |
| CM | 3532 | 5925 | 7378 | 9927 | 26852 |
| Sent Letter (IV) | 149.02 | 267.11 | 330.4 | 546.06* | 1348.83** |
| | (106.66) | (177.76) | (238.80) | (317.24) | (685.56) |
| CCM | 3682 | 6132 | 7554 | 10239 | 27661 |

Notes: N = 43,409. Panel B shows raw earnings with a single outlier (>$3 million in earnings in one quarter) top-coded to next highest earnings in data. Winsorization in Panel C recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings in a given year before summing to yearly totals. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 3: Amount and Timing of Work

| | Num Quarters Worked | Num of Job Spells | Num of Job Spells if >0 | Avg Spell Length (Spell 1-3) | Time to First Qtr Worked |
|---|---|---|---|---|---|
| ITT | 0.062 | 0.019 | 0.010 | 0.047** | -0.103*** |
| | (0.041) | (0.022) | (0.022) | (0.024) | (0.030) |
| CM | 7.28 | 3.43 | 3.72 | 3.13 | 3.01 |
| Sent Letter (IV) | 0.152 | 0.046 | 0.025 | 0.117** | -0.254*** |
| | (0.101) | (0.054) | (0.054) | (0.058) | (0.074) |
| CCM | 7.53 | 3.43 | 3.71 | 3.24 | 3.01 |
| N | 43409 | 43409 | 40088 | 40088 | 40088 |

Notes: Spells are defined as consecutive quarters with earnings from same employer. The third through fifth columns condition on having at least one spell. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 4: Job Application Effects

| | Clicked Link | Applied | Checked Selective Job Box | Uploaded Any File | Included Letter of Rec |
|---|---|---|---|---|---|
| ITT | -0.007 | -0.006 | -0.010 | 0.003 | 0.010*** |
| | (0.009) | (0.009) | (0.007) | (0.007) | (0.003) |
| CM | 0.103 | 0.088 | 0.053 | 0.052 | 0.004 |
| Sent Letter (IV) | -0.020 | -0.019 | -0.027 | 0.006 | 0.024*** |
| | (0.024) | (0.022) | (0.017) | (0.018) | (0.007) |
| CCM | 0.138 | 0.123 | 0.082 | 0.065 | 0.009 |

Notes: N = 4,000. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 5: Relationship between Ratings and Other Outcomes among Controls

| Dependent Variable: | Cumulative Earnings | | GPA Year 1 | | Graduated or Still Attending | |
|---|---|---|---|---|---|---|
| | Panel A: Full Sample Control Group | | | | | |
| Rating | 1918.67*** | 971.70*** | | | | |
| | (171.17) | (153.14) | | | | |
| Mean | 27,243 | | | | | |
| N | 15487 | | | | | |
| | Panel B: Education Sample Control Group | | | | | |
| Rating | 792.77*** | 607.27*** | 1.80*** | 0.30*** | 0.027*** | 0.005** |
| | (170.94) | (164.50) | (0.10) | (0.06) | (0.003) | (0.002) |
| Mean | 18,394 | | 80.14 | | 0.855 | |
| N | 7053 | | 6532 | | 7053 | |
| Covariates | No | Yes | No | Yes | No | Yes |

Notes: Coefficients from regressing each dependent variable on supervisor rating in the control group. Earnings and school persistence measured across 4 post-randomization years. Regressions include every control individual with a non-missing supervisor rating who is part of the the main sample (Panel A) or education sample (Panel B). Columns marked as having covariates include all the available baseline covariates for each sample listed in Appendix Section A.5. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 6: Labor Market Effects for Youth with High and Low Supervisor Ratings

| | | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | Employment | | | | |
| ITT, Low Ratings | | 0.0247* | -0.015 | -0.0169 | 0.0096 | 0.0001 |
| | | (0.0133) | (0.0134) | (0.0139) | (0.0141) | (0.0080) |
| ITT, High Ratings | | 0.013 | 0.0237*** | 0.0157* | 0.0067 | 0.0088* |
| | | (0.0087) | (0.0087) | (0.0092) | (0.0093) | (0.0052) |
| P-value, test of diff. | | 0.463 | 0.015 | 0.051 | 0.865 | 0.361 |
| CM, Low | | 0.673 | 0.721 | 0.657 | 0.669 | 0.924 |
| CM, High | | 0.715 | 0.720 | 0.656 | 0.687 | 0.925 |
| | First Stage | | | | | |
| IV, Low Ratings | 0.3301*** | 0.0747* | -0.0454 | -0.0511 | 0.0291 | 0.0002 |
| | (0.0103) | (0.0404) | (0.0405) | (0.0422) | (0.0428) | (0.0242) |
| IV, High Ratings | 0.8108*** | 0.0161 | 0.0292*** | 0.0194* | 0.0083 | 0.0107* |
| | (0.0057) | (0.0108) | (0.0108) | (0.0114) | (0.0115) | (0.0064) |
| P-value, test of diff. | 0.000 | 0.161 | 0.075 | 0.107 | 0.639 | 0.675 |
| CCM, Low | | 0.613 | 0.757 | 0.688 | 0.668 | 0.917 |
| CCM, High | | 0.713 | 0.717 | 0.657 | 0.687 | 0.924 |
| | | Earnings, Winsorized at 99th Percentile | | | | |
| ITT, Low Ratings | | 25.60 | -186.00 | -191.12 | -58.83 | -355.69 |
| | | (121.61) | (207.51) | (266.56) | (347.72) | (758.53) |
| ITT, High Ratings | | 92.02 | 323.69** | 222.78 | 562.34* | 1239.98* |
| | | (92.97) | (161.62) | (219.37) | (295.36) | (633.15) |
| P-value, test of diff. | | 0.664 | 0.053 | 0.231 | 0.173 | 0.106 |
| CM, Low | | 3083 | 5369 | 6474 | 8361 | 23323 |
| CM, High | | 3679 | 6205 | 7918 | 10880 | 28811 |
| | First Stage | | | | | |
| IV, Low Ratings | 0.3301*** | 75.62 | -575.42 | -586.89 | -184.04 | -1099.11 |
| | (0.0103) | (368.07) | (628.68) | (807.87) | (1053.00) | (2298.22) |
| IV, High Ratings | 0.8108*** | 112.86 | 394.95** | 266.12 | 687.63* | 1517.59* |
| | (0.0057) | (114.66) | (199.31) | (270.62) | (364.26) | (780.87) |
| P-value, test of diff. | 0.000 | 0.923 | 0.141 | 0.317 | 0.434 | 0.281 |
| CCM, Low | | 3115 | 5976 | 7047 | 8891 | 24887 |
| CCM, High | | 3589 | 6045 | 7830 | 10785 | 28342 |

Notes: To avoid selection into who is rated within a survey, sample includes only youth on a survey where the supervisor rated all listed youth (N = 13,911). Low = rating categories 1–4; High Ratings = rating categories 5–7. P-value from tests of the null hypothesis that treatment effects are equal in low and high ratings groups. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table 7: Education Effects

| | Ever Enrolled Y1 | % Enrolled Days Present Y1 | GPA Y1 | Credits Attempted Y1-4 | % Credits Earned Y1-4 | Graduated On-Time | Ever Graduated | Graduated or Still Attending | On-time College |
|---|---|---|---|---|---|---|---|---|---|
| ITT | -0.002 | 0.001 | -0.130 | 0.063 | 0.003 | -0.007* | -0.001 | -0.004 | -0.004 |
| | (0.003) | (0.003) | (0.099) | (0.100) | (0.003) | (0.004) | (0.004) | (0.004) | (0.006) |
| CM | 0.946 | 0.829 | 80.13 | 18.96 | 0.818 | 0.785 | 0.834 | 0.851 | 0.672 |
| IV | -0.005 | 0.003 | -0.302 | 0.194 | 0.006 | -0.016* | -0.002 | -0.009 | -0.010 |
| | (0.007) | (0.007) | (0.237) | (0.241) | (0.007) | (0.010) | (0.010) | (0.010) | (0.013) |
| CCM | 0.957 | 0.849 | 81.75 | 18.49 | 0.844 | 0.828 | 0.865 | 0.883 | 0.717 |
| N | 19714 | 19714 | 18237 | 19714 | 19714 | 19714 | 19714 | 19714 | 17810 |

Notes: Analysis is conducted on all those expected to be observed based on pre-program grade of enrollment (see text for details). Percent credits earned is number earned divided by number attempted. Credits attempted and percent credits earned equal 0 for those not in school. On-time graduation equals 1 for public school, non-charter students who graduate within 4 years and do not transfer to GED programs or other districts. Ever graduated adds any 5th- and 6th-year graduation observed during the follow-up period. Graduated or still attending equals 1 if student either graduated or has positive days attended in most recent academic year. College enrollment is only measured within 6 months after a student's on-time graduation date, regardless of graduation status. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are shown in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Figure 1: Experimental Flow Chart

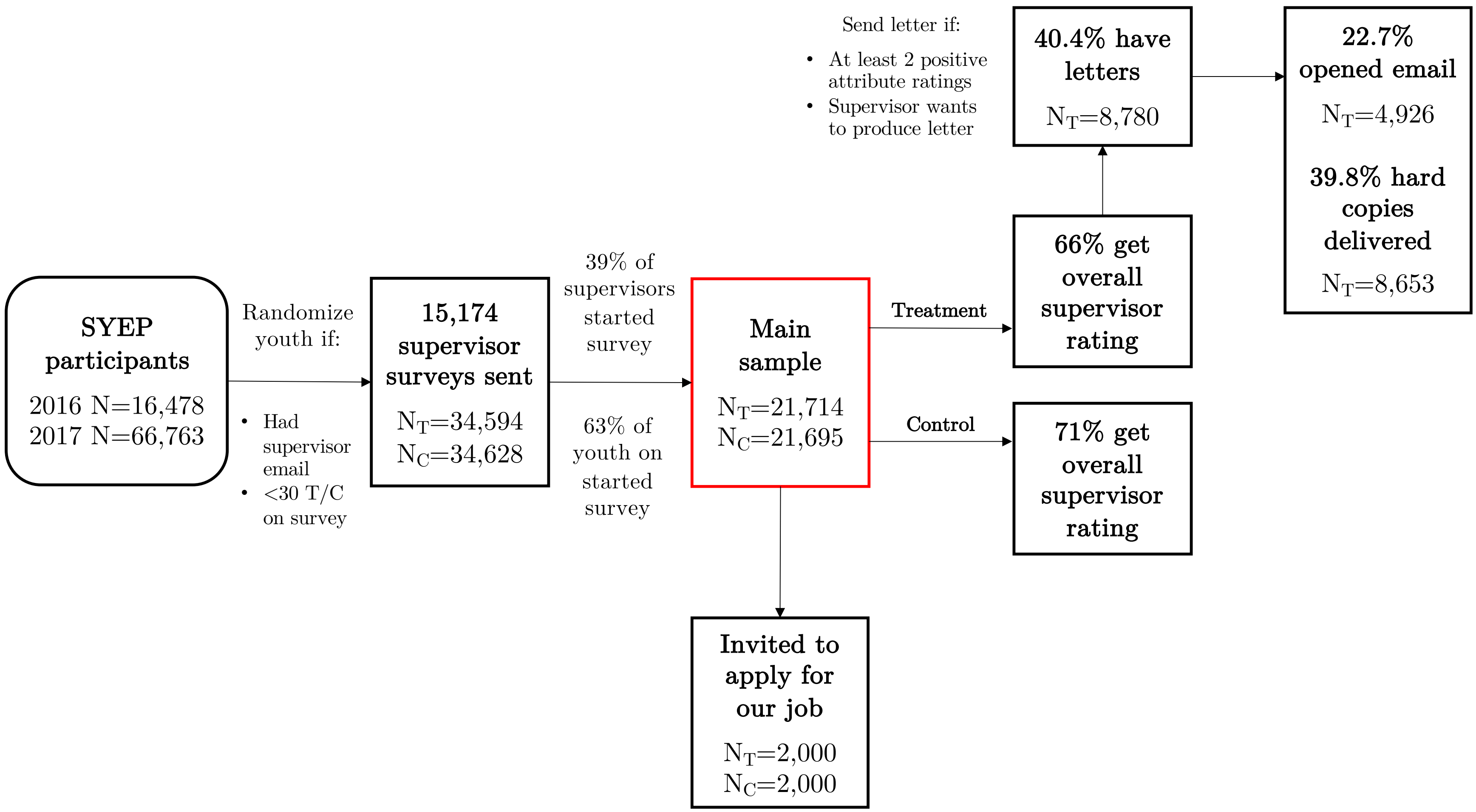
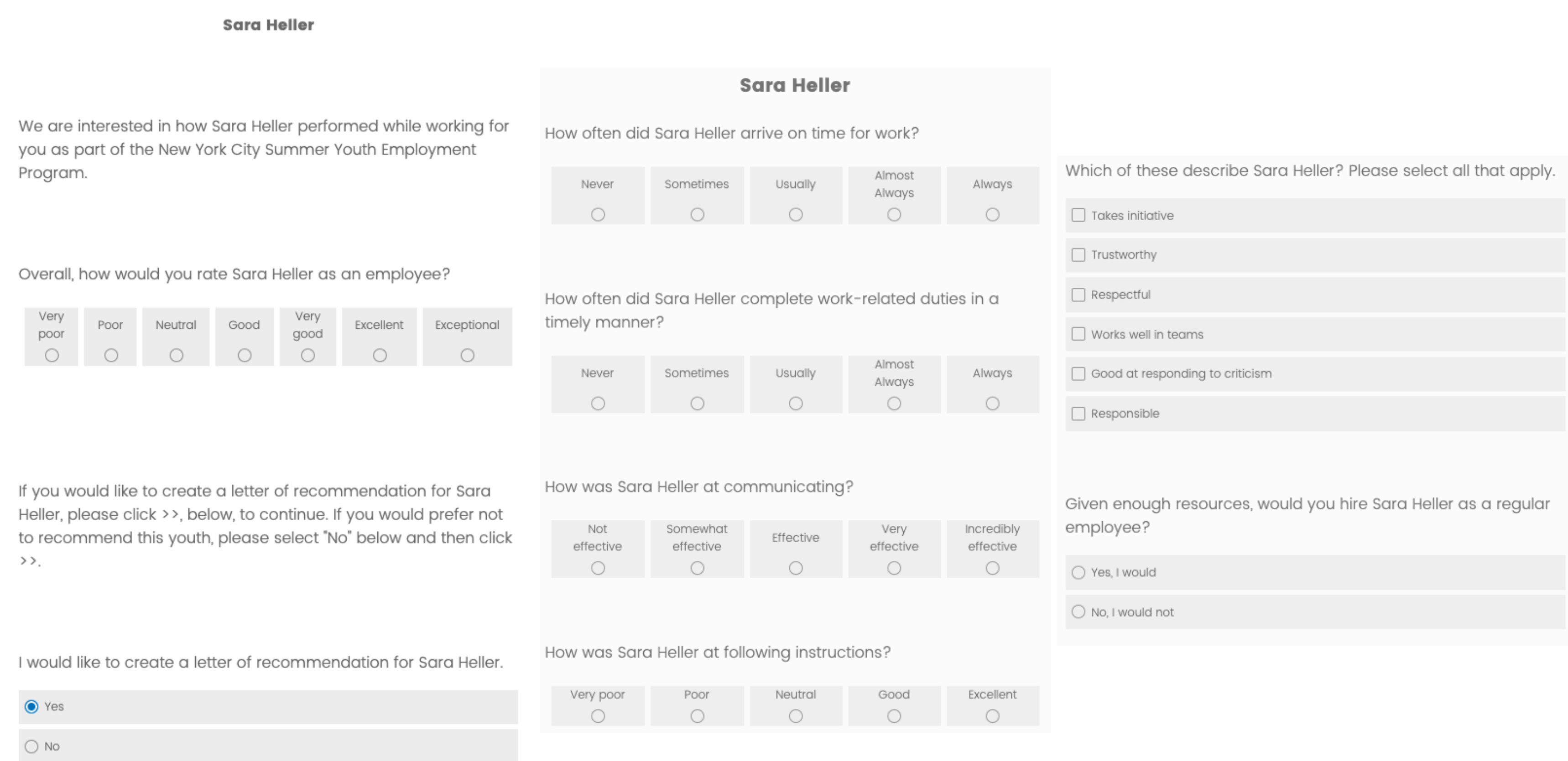


Figure 2: Screenshots about Treatment Youth on Supervisor Survey




Notes: The image on the left shows the first screen supervisors saw asking about each youth with the overall rating question and the invitation to write a letter. As indicated in the image, the option to create a recommendation was pre-selected. The images in the middle and on the right show the questions asked about each treatment youth when the supervisor agreed to create a letter of recommendation.

Figure 3: Example Letter of Recommendation

Department of
Youth & Community
Development

November 1, 2017

To Whom It May Concern:

Sara Heller worked for me at the Wharton School during the summer of 2017. Overall, Sara was an exceptional employee.

With regard to reliability, Sara was always on time to work. Sara always completed work related tasks in a timely manner.

When it came to interpersonal interaction, Sara was an incredibly effective communicator. Sara was excellent at following instructions.

In addition to Sara's other strengths, Sara takes initiative, is trustworthy, is respectful, works well in teams, is good at responding to constructive criticism, and is responsible.

Given the resources, I would hire Sara as a full-time employee. I invite you to contact me if you would like more information. I can be reached at 215-898-7696 or judd.kessler@wharton.upenn.edu.

Sincerely,

Judd Kessler
The Wharton School

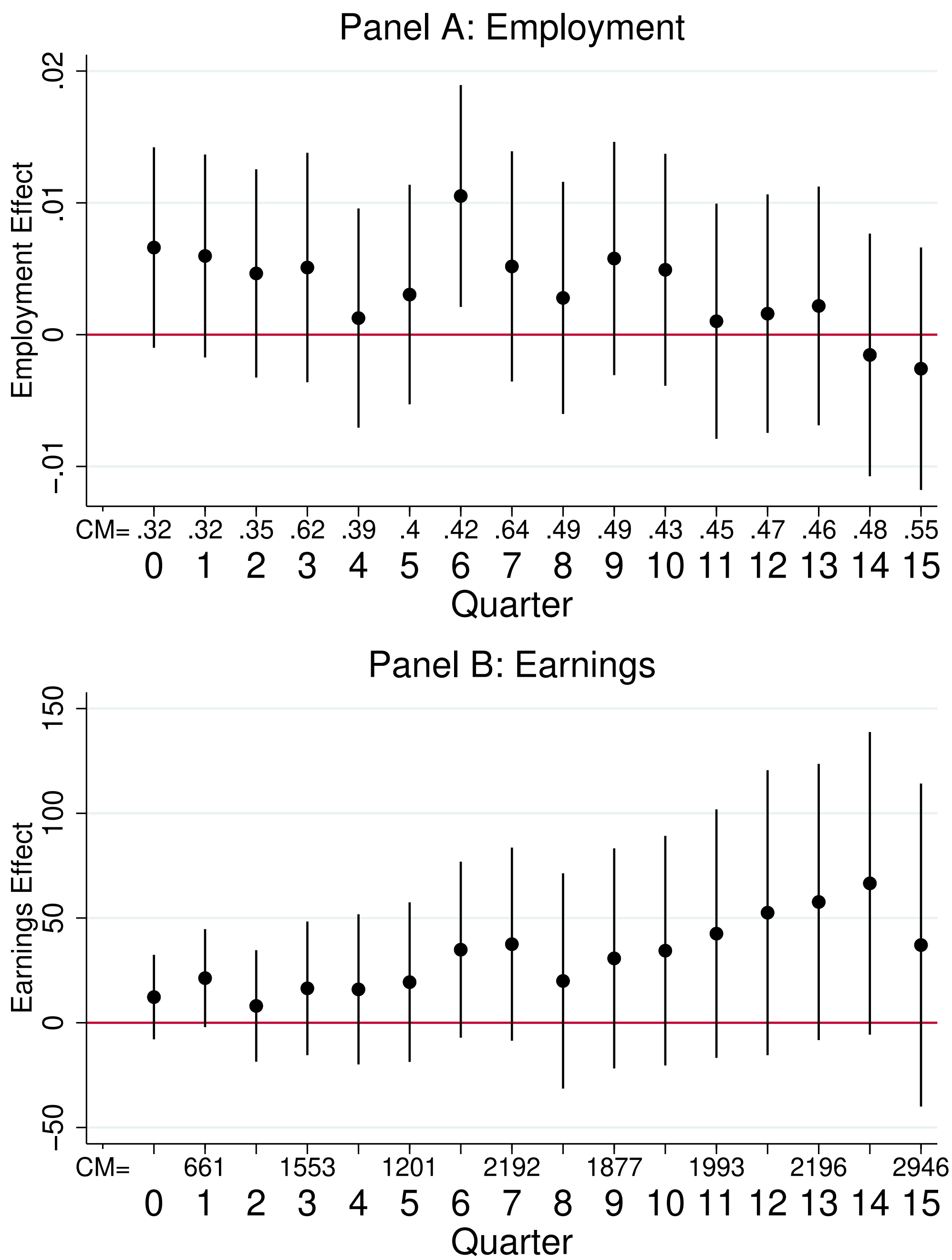
The New York City Department of Youth and Community Development (DYCD) invests in a network of community-based organizations and programs to alleviate the effects of poverty and to provide opportunities for New Yorkers and communities to flourish.

Empowering Individuals • Strengthening Families • Investing in Communities

Note: This recommendation letter is part of a pilot program being run by the New York City Department of Youth and Community Development. Some youth were randomly selected to be part of the pilot. These youth were eligible to receive a letter of recommendation, which reflects supervisor feedback about each individual's job performance.

Notes: See Figure 2 for the source of inputs into each sentence for this example letter.

Figure 4: Labor Market Effects by Quarter




Notes: Figure shows intent-to-treat effects on employment and earnings (winsorized at the 99th percentile) by quarter, with 95 percent confidence intervals calculated from standard errors clustered on individual. CM below axis displays the control mean in that quarter. All effects from regressions that include baseline covariates. Letters were distributed in Quarter 0. Quarters 0–3 comprise year 1, 4–7 are year 2, 8–11 are year 3, and 12-15 are year 4. Summers (July–September) are quarters 3, 7, 11, and 15.

# ONLINE APPENDIX

# Information Frictions and Skill Signaling in the Youth Labor Market

Sara B. Heller
University of Michigan & NBER

Judd B. Kessler
The Wharton School & NBER

## A.1 Additional Labor Market Results

### A.1.1 Earnings

The main text shows that letters of recommendation increase employment in the short term and earnings in the longer term. A natural question is whether the earnings increase comes from additional part-time employment or from shifting people into high-paying or full-time work. Although we can not observe hours to know for sure, we can look at the full earnings distribution by treatment group to get a sense for where the shifts in the distribution occur. Panel A of Figure A.8 shows the full raw earnings distribution (with the single extreme outlier top-coded, as in the main text). Because the treatment effects are small relative to the scale of all earnings over 4 years, it is hard to see them in Panel A (although it is clear that there is a control outlier, which helps to explain the sensitivity to different skewness adjustments discussed below).

Panel B of Figure A.8 zooms in on the bottom of the distribution, under $25,000 in 4-year earnings, to make the treatment effect more visible (the distributions above $25,000 are quite similar). This figure suggests that the bulk of the treatment effect comes from moving people near zero up to earning between $2,000 and $5,000. Over 4 years, this pattern is most easily explained by treatment youth having an additional part-time job. Figure 4 in the main text shows that this change was not just over the summer; employment and earnings effects are similar in summer quarters and other quarters.

There is also a smaller shift away from earnings between $5,000 and $8,000 and into earnings between $20,000 and $24,000. These higher earnings are consistent with more persistent part-time work, or potentially a year of minimum wage, full-time work. Overall, the earnings distributions suggest that the letters help people enter the labor force or shift to somewhat more or better-paying work; they are not shifting youth into high-paying, full-time jobs.

The main text reports annual and cumulative earnings results for two functional forms of the earnings variable: raw (with one extreme outlier observation top-coded) and winsorized at the 99th percentile. Because we pre-specified that we would explore other adjustments for skewness, Table A.1 shows other transformations of the raw dollar amounts, including an alternative winsorization (at the 99.5th percentile), log earnings with different intercepts added to assess how much the infinite proportional change from 0 matters [log(earnings + 0.1, 1, 10, or 100)], and the inverse hyperbolic sine transformation.

The alternative winsorization in Panel A makes very little difference relative to the results in the main text. The other panels show that, as expected given that there are treatment effects on the extensive margin, the decision about what to add to the 0s does change the point estimates somewhat. Because the biggest change on the employment margin is in year 1, results in year 1 are most sensitive to what is added to 0. The results range from a 9.5 percent increase to a 30 percent increase in year 1 earnings, driven by the fact that so many people are moved off of 0, where the proportional change is undefined. Since fewer people are moved off of 0 for the cumulative earnings measures, those results are more sensible in magnitude, ranging from a 7 percent to a 12 percent increase in earnings over the four years. We emphasize the 4.9 percent increase in the main text, both because the winsorized results were our primary pre-specified outcome and because it is clear that the logged results are sensitive to how we handle the 0s.

### A.1.2 Spell Length

The fourth column of Table 3 in the main text shows that treatment increases the average spell length among the first 3 (non-missing) spells. We argue that this result is an indication of improved job match quality among treatment youth relative to control youth. One reason to care about this result is that it pushes against the hypothesis that letters drive employers to inefficiently update (e.g., as might happen if previous applicants with letters were always stellar employees and employers incorrectly believe that any applicant with a letter will be similarly stellar). If employers did inefficiently update in this way, we might expect them to be more willing to hire treatment youth, but then to quickly fire them after learning that they were not as high productivity as expected, which would create inefficient churn. The fact that spell length increases with treatment, however, suggests that letters' signals instead help employers to successfully identify good matches.

The main text reports that across the 3 spells underlying Table 3, there is no treatment-control difference in the number of censored spells, despite treatment spells starting earlier. Table A.2 provides additional evidence on this pattern by looking separately at each of these spells. Each panel shows results for a different job spell, with spell 1 being the spell started the earliest, spell 2 being the spell started next, and so on. If spells are started in the same quarter, we assign the longer spell the lower spell number. We count any spell with at least one quarter occurring in the post-letter period. Youth must have a given spell number to appear in each panel, so the sample becomes more selected as the spell number rises (about 60 percent of the sample has a third spell). The first column reports treatment effects on the length of each spell, defined as the number of consecutive quarters worked at the same employer. The treatment effect on the length of individual spells is always positive, but imprecisely estimated when broken down by individual spell.

The control means suggest why differential censoring may not be a problem for these early spells: even the earliest spell has an average length of just under 4 quarters, so only 7 percent of them are censored (defined as a youth working at an employer in the last quarter we observe in the data). Censoring rises to about a quarter of third spells. We stop at spell 3 to avoid too much further censoring, and because the average number of spells in the sample is just over 3.

As shown in the second column, none of the censoring is significantly different by treatment group, suggesting that differential censoring is not biasing our spell length results. The last 3 columns of the table confirm that the results are robust to looking only at spells that are not censored. We report treatment effects on whether a spell lasts at least 2, 3, or 4 quarters, conditional on observing all the quarters. There is no evidence that letters are creating bad matches, with all but one of the point estimates positive. Overall, analysis at the individual spell level is a bit imprecise, which leads us to average these spell lengths (and report the censoring result across all 3 spells) in the main text.

### A.1.3 Employer Type

Tables A.3 and A.4 separate employment and earnings effects by type of employer. Because the letter came on DYCD letterhead (the agency that runs the SYEP), it is possible that the letter increased the rate at which youth reapplied to the SYEP or engaged with future summer or term-time work where DYCD was the employer of record.

Table A.3 shows that this is not a main driver of our results. It reports labor market results separately for DYCD and for all other employers. The only significant increase in employment is at non-DYCD employers, meaning that the letters increased employment outside of the SYEP agency. Earnings impacts are directionally much larger at non-DYCD employers, on the order of 5 rather than 1 percent.

Table A.4 shows in what types of industries letter recipients work. The classification across industry clusters is based on Gelber, Isen, and Kessler (2016), which groups industries that are over-represented in SYEP, like childcare and landscaping (cluster 1) separately from industries that are under-represented in SYEP, such as retail and food service (cluster 2). Letters directionally increase employment in both types of industries, only marginally significant in year 1 for cluster 2, with earnings increases concentrated in cluster 2 jobs. This pattern suggests that the letters are helping young people shift to jobs outside of the industries that they were most likely to be exposed to through SYEP. Given the evidence from Gelber, Isen, and Kessler (2016), which found that working in cluster 1 jobs results in lower overall earnings than the cluster 2 jobs, the patterns here are consistent with treatment youth using their letters to shift towards higher-paying industries.

## A.2 Supervisor Ratings

Panel A of Figure A.9 shows the overall distribution of ratings that supervisors assigned to both treatment and control youth. As discussed in the main text, we designed the survey to maximize the information we would have available to produce recommendation letters, not to ensure that treatment and control youth would be treated equally on the survey. As such, we asked about each treatment youth first, on the same page as we asked supervisors to decide whether to produce a letter. After the supervisor had seen all treated youth, we then asked them a single question about the overall performance of each control youth—all on the same page—making it clear the control youth were not eligible for letters. This aspect of our design makes it possible that supervisors might use different decision rules across treatment and control youth when assessing whether to give a rating and what rating to give.

Indeed, treatment youth are significantly less likely to have been rated by a supervisor (66 versus 71 percent had a rating, $p<0.001$). Panel B of Figure A.9 shows that treatment youth have a more compressed ratings distribution, with missing mass on both the highest and lower rating categories. This pattern might indicate that supervisors take the letters seriously, so are less likely to give very top marks when they know their responses will be included in a letter, but also less likely to give someone the lowest marks (perhaps to be kind to the youth, since supervisors did not know our exact decision rule for when not to send a bad letter).

Despite the potential for selection into having a rating, observable characteristics are generally still balanced in the sample with non-missing ratings, with a joint F-test (including the actual rating) failing to reject equality across all observables ($p = 0.609$). Table A.5, however, which breaks out the balance tests for youth receiving low versus high ratings, shows that there is some imbalance within the group that receives low ratings ($p = 0.101$). Since breaking out the results by rating group is central to understanding whether employers are using the letters as signals to accurately update their beliefs, the potential for selection within the rating groups is of concern.

Because of the dramatic difference in having a rating and the small imbalance on observables for those with low ratings, the main text focuses on the subsample of rated youth on complete surveys. Table A.6 shows the equivalent balance tests for the subsample of youth who appeared on a fully completed survey (i.e., where the employer rated every youth on the survey). Although this is a selected group, full survey completion limits the scope for treatment and control youth to be differentially selected into getting a rating. Indeed, the difference in receiving a rating is much smaller in this sample: 31.6 percent for treatment youth and 32.5 percent for controls ($p = 0.066$). And, as the table shows, observables are entirely balanced within each rating group. (Panel C of Figure A.9 suggests there may still be some differences in exact ratings, but only across the ratings that are all classified as "high" in our regressions.) As a result, this is the subsample we use to assess how treatment effects vary by rating in the main text.

Despite our concern about the potential for selection, for completeness, Table A.7 shows the main labor market effects for everyone with a rating, without limiting the sample to completed surveys as in the main text. The patterns are fairly similar to the results in the main text, with the high-rated group showing significantly positive employment effects, especially in the early years, and much more positive earnings impacts than the low-rated group. The earnings point estimates are a bit smaller than in the main text and so not statistically significant outside of year 1, though they still generally grow over time for the high-rated group. In this sample, the low-rated group (where there is the most observable imbalance) has somewhat more positive employment effects, but still has negative earnings point estimates.

We have also tested whether treatment effects on applying to our job posting are different for those with a high versus low rating. Given that this limits an already reduced sample (N = 4,000) to those with ratings (N = 2,783, when we use all ratings), and then splits the sample into groups, this is not a highly powered test. The difference in the intent-to-treat effects for the high-rated group relative to the low-rated group (i.e., the interaction effect between treatment and being highly rated) is $\beta = 0.008$, $p = 0.721$, with a control rate of application for the low group of 0.078. The difference for the IV is $\beta = 0.018$, $p = 0.748$. So while it is possible that receiving a letter had a more positive effect on job search behavior for highly-rated youth, we cannot reject the null that both effects were zero.

## A.3 Heterogeneity

Tables A.8 through A.15 show treatment effects for different subgroups of youth. Because of the number of hypothesis tests across these tables and the limited statistical power, we do not emphasize the statistical significance of any particular result. However, we pre-specified an interest in these divisions as exploratory, so we report the basic patterns here.[1a]

### A.3.1 Mechanisms and Heterogeneity

It is tempting to use basic cuts of the data to help understand the mechanisms driving our main effects. But theory makes clear that single cuts of the data may not be enough. Consider the prediction from the statistical discrimination literature that those with fewer available signals should benefit more from a new signal. That might tempt us to interpret heterogeneity by whether someone ever worked, for example, as a test of statistical discrimination, if we think having no work history means there is more uncertainty about performance.

Importantly, however, as Pallais (2014) proves, theoretical predictions about heterogeneity for these groups are not clear cut. It is only *conditional on ability* that signals should have a bigger effect for those with more uncertainty. If those without signals (e.g., those who have never worked) also have lower average productivity, it is not evident that signals should help that group more. If the letters more often reveal that those with no work history are less prepared for work, we should not expect the signal to improve labor market success.

Given our setting, there are a number of other factors that also vary by subgroup: whether supervisors generate a letter, how strong the letter is, whether workers are looking for work, and whether they decide to use a letter in their applications. To help interpret our subgroup effects, we report the first stage by group, and we summarize application and letter use behavior by group in Table A.16. That said, we emphasize that the many different factors that vary by subgroup make it hard to convert treatment heterogeneity into a clear mechanism story. Doing so would likely require significant assumptions about the structure of the job search process. In addition, as our pre-analysis plan anticipated, we are not well-powered for heterogeneity tests. As a result, while we report subgroup effects—to aid in comparisons to prior work and because descriptive patterns of subgroup results help speak to general questions about labor market inequality—we are cautious not to over-interpret

1a. We add two divisions that were not pre-specified: whether someone is in our education sample and whether they had worked prior to the summer of the SYEP. The former both helps to check whether labor market effects differ for the sample underlying the main education results and provides a rough cut by whether individuals are still in high school (though some of our sample is in high school but not in our education sample, because, e.g., they attend schools that are not in our education data).

these patterns.

### A.3.2 Heterogeneity by Subgroups

Table A.8 compares labor market impacts for those who are and are not in the expected in high school sample. Both groups respond positively to the letters. The employment effects are slightly more persistent for those in the education sample, though cumulative earnings impacts are almost identical.

Table A.9 shows effects for those under 18 and those 18 and over at the time of application. Employment point estimates are slightly larger and earnings estimates slightly smaller for those under 18, but both sets of effects are statistically indistinguishable from the effects for older youth.

Table A.10 shows labor market impacts separately for young people who did or did not have any prior work experience (measured as appearing in the UI data) before the SYEP summer. Point estimates are larger and only statistically significantly different from zero for the group that had previous work experience, which is a similar finding as in Pallais (2014). This result is perhaps more consistent with the possibility that employers are using the letters to help identify those likely to be higher performers, rather than to just improve their priors about those with the least available information.

Table A.11 shows results separately for White and non-White youth. The latter group includes youth who are Black, Hispanic, Asian, and Mixed Race/Other in the SYEP data. All the main labor market effects are concentrated among non-White youth, with cumulative earnings effects marginally different from each other.

Tables A.12 and A.13 further break down the main labor market results separately by race and ethnicity subcategories (ITT and IV, respectively). They show that the employment impact is driven by somewhat larger effects for Asian and Hispanic youth, and to a lesser extent those in the Other category, with earnings effects suggestively larger as well. The likelihood of getting a letter is higher for these groups than for Whites (see first column of Table A.13), but even among compliers, the program impacts are larger for Asian, Hispanic, and Other youth. However, as in the main results, we are under-powered to detect group differences; we cannot reject the null that effects are the same across all groups.

Table A.14 shows that female SYEP participants are significantly more likely to receive a letter, with female compliers having suggestively larger employment effects in year 1. In contrast, earnings effects are quite similar by gender; if anything, men have slightly larger point estimates for earnings. The initially larger employment effect for women is consistent with the Abel, Burger, and Piraino (2020) result that the employment benefits of recommendation letters in South Africa were concentrated among women. But unlike in that setting, young women in NYC do not face the same difficulty finding work relative to young men; indeed, consistent with broader U.S. patterns of young women outperforming their male counterparts, employment rates for women are considerably higher than for men in our sample. The fact that there are larger effects for women both in settings where priors are likely to favor and to disfavor women suggests that the effect is not simply about statistical discrimination, since priors should go in the opposite direction across settings. Additionally, our longer-term results suggest overall effects are fairly similar across gender.

Table A.15 shows effects by neighborhood economic mobility. Using the Opportunity Insights "upward mobility" data (https://opportunityinsights.org/data/), we use each individual's zip code to assign their neighborhood an average income rank for children whose parents were in the 25th percentile of the national household income distribution. Opportunity Insights provides these data at the Census Tract level. We use the Zip Code Tabulation Area (ZCTA) crosswalk to map Census Tracts onto zip codes, which is the geographic information we have on our sample. In cases of multiple Census Tracts falling within a given ZCTA, we use the average upward mobility value (i.e., the unweighted mean across all upward mobility values that fall within the ZCTA). We divide the youth into those who live in areas with above and below median mobility, with median defined in-sample. Table A.15 shows labor market impacts for these two groups. There are positive effects for both those living in above-median and below-median neighborhoods, with early employment effects suggestively larger in places with below-median mobility, but earnings effects suggestively larger in places with above-median mobility.

### A.3.3 Information on Letters by Subgroup

To help interpret the patterns of results by subgroup, Table A.16 shows some additional information about the letters for the different subgroups discussed in the previous section. The table shows the treatment group only, since they were the only ones eligible for a letter. The first column shows the proportion of each group that was sent a letter (i.e., having a supervisor agree to produce one and receiving ratings high enough to generate a letter); this summarizes the information shown in the "first stage" column of the separate heterogeneity results. The second column is conditional on the first, showing average overall employee rating on a scale from 1–7 for those who were sent a letter. The third column shows the proportion of each group that submitted an application in response to our job application, conditional on being one of the 2,000 treatment youth randomly selected to receive the job advertisement. The fourth column, conditional on the third, shows the proportion of the applicants that uploaded a letter of recommendation (ours or any other) as part of their application.

There is significant variation both in letter receipt and in average ratings. Non-white, female, in high school, previously-employed, and below-median neighborhood mobility youth are all more likely to receive a letter. But the higher rate of letter receipt does not always correspond with stronger letters, on average. For example, despite larger labor market impacts, non-White youth have significantly lower average ratings conditional on receiving a letter than their White counterparts. And they do not use the letter more frequently; their rate of letter usage is about 6 percentage points lower than the White youth who applied to our job posting, although the small sample size limits how well we can differentiate the groups. The basic pattern of results suggests that the larger labor market effects for non-White youth are likely to be driven by how employers respond, even to slightly weaker letters, rather than big differences in how the groups use the letters.

The only significant differences in letter usage are between those who were or were not in our education sample at the time of SYEP application, and relatedly, those who were under 18 versus 18 and older. This likely helps to explain the bigger employment point estimates for our education sample, who were much more likely to use the letter on our job application than those who were not expected in our school data.

# A.4 Additional Education Results

## A.4.1 Descriptive Statistics

Table A.17 shows descriptive statistics and treatment-control balance for our education sample. On average, students in our education sample are about 16 years old, 45 percent male, 42 percent Black, 31 percent Hispanic, 14 percent Asian, and 8 percent White. They are in 10th grade on average, attending about 90 percent of the days they are enrolled, and earning a C-plus average. Over 60 percent of them had not worked in UI-covered jobs prior to the SYEP. The table also shows that across all baseline characteristics, treatment and control groups are jointly balanced ($p = 0.149$). It is worth noting that there is some chance imbalance on GPA and on the proportion of the sample that is White; although the differences are substantively small (-0.39 on a 100-point GPA scale and 1 percentage point more likely to be White), they are statistically significant. As a result, the exact magnitude of the education results are somewhat more sensitive to how covariates are included in the regressions (see Appendix Section A.5). However, none of our substantive conclusions are sensitive to covariate choice.

### A.4.2 Joint Work and Graduation Outcomes

In the main text, we note that there is evidence that the decrease in on-time graduation is driven by the same youth who are pulled into the labor force. This claim comes from examining the relationship between educational attainment and labor force involvement within the same individual. We define a set of mutually exclusive joint outcome indicators: working and graduating, never working but graduating, working and not graduating, and never working and not graduating. We define these indicators for all three of our education attainment measures: on-time graduation, ever graduating, and graduating or still attending school.

The treatment effects across these outcomes allow us to assess whether any potential shifts in educational attainment occur among the same group that experiences shifts in employment. Table A.18 shows the results. The third column of Panel A shows that there is a significant increase in the proportion of people who work but do not graduate on time of about 2.3 percentage points (16.6 percent) for compliers. Since everyone has to appear in one and only one of the columns, the other columns' estimates show where the marginal work-but-not-graduate-on-time group comes from. The shift to the third column appears to be spread across the other categories, with the biggest shifts from reductions in the number of people who both work and graduate on time, as well as those who neither work nor graduate on time. Although the results in the other columns are not significant, the point estimates suggest that some of those shifted by the letter just add work on top of what would have already been a failure to graduate on time. But for others, the letters seem to prevent them earning their on-time diploma.

Panel B, which measures whether people ever graduate and work, suggests that the decline in on-time graduation may not be permanent. There are no significant changes in work/ever graduate categories. The point estimate for working but not graduating is about a third as large as in Panel A, and there is also a positive point estimate for both working and graduating. The combination of Panels A and B is what drives our conclusion in the main paper that it is the shift into the labor force that slows down graduation, but that it appears most of the slowed-down students will eventually graduate.

Panel C provides some caution, though. By including continued school attendance as part of the dependent variable, it aims to capture what happens to students who are either too old to show up in the graduation data (graduating after their 6-year cohort) or too young to have reached their final graduation outcome. It suggests that there is still a letter-driven increase in working but not persisting in school. While about half of this shift appears to come from people who would otherwise not have worked or graduated (as indicated by the negative point estimate in column 4), the other half seems to shift from groups that would otherwise have persisted (columns 1 and 2). In combination with Panel B, this might suggest that at least some of the students who could eventually graduate are not still attending school. Longer-term follow-up is needed to assess what these students' final outcomes will be; it is not uncommon for people at the margin of graduating to leave school temporarily and return later.

Nonetheless, these results suggest some caution about encouraging youth at the margin of school completion to join the labor force. The following section further explores this margin by splitting students by baseline academic achievement.

### A.4.3 Explaining the Decline in On-time Graduation: Heterogeneity by GPA

In the main text, it is not entirely clear how seriously to take the marginal decline in on-time graduation, since no other educational outcomes show significant declines. If letters are truly slowing down graduation, we might expect to see the mechanisms through which that happens in some of our educational performance measures. In this section, we assess whether there is real concern that the increase in labor market participation prevents a subgroup of youth from the educational progress they would otherwise make. We do this by examining heterogeneity that should be closely related to whether youth are on the margin of graduation: baseline GPA. We split the education sample by whether students are over or under the median GPA in the baseline year (for non-missing GPAs only, n = 17,732, median GPA = 80.85).

Table A.19 shows the main education outcomes by GPA, focusing on the IV to conserve space, and Table A.20 shows the corresponding labor market outcomes, including the first stage. Above-median GPA students show no significant changes in education outcomes. But the top row of Table A.19 demonstrates that letters do, in fact, harm the educational progress of the below-median GPA students. They have lower year 1 enrollment (by 2.5 percentage points, or 2.6 percent), perhaps indicating that receiving a letter in the fall of the academic year deters some students from returning to school the following semester. Those that remain in school have significantly lower GPAs (by 0.85 points on a 100 point scale, or 1.2 percent). And though the increase in credits attempted is not statistically significant, it is positive, suggesting some of the drop in GPA might result in retaking courses, which could slow down graduation. Indeed, the decline in on-time graduation is larger and more statistically significant in this subgroup (5.7 percentage points, or 7.6 percent).

As in the main sample, the point estimate on whether below-GPA students ever graduate is considerably smaller than for on-time graduation (-0.02 compared to -0.06), suggesting that at least some of those who are delayed catch up and eventually graduate. But overall school persistence and on-time college enrollment also have negative point estimates, so final conclusions may need to wait until everyone has had time to either graduate or leave school more permanently.

Consistent with the idea that it is increased labor force participation driving the educational changes, Table A.20 shows that the below-median students have a significantly larger increase in employment in year 1 that remains substantively large but not significant in years 2 and 3, with a significant increase on the intensive margin of work (number of quarters worked) in year 2. Above-median GPA students still benefit from letters, but largely with higher earnings rather than more employment. Table A.21 confirms that the changes in joint outcomes are also concentrated among the below-median students, including declines in persistence. So the bigger boost into the labor market appears likely to be pulling these marginal students out of school.

From a policy perspective, these results provide some caution against the recent push for governments to offer year-round work opportunities to students who might not otherwise obtain term-time jobs. Contrary to results using natural variation in work during school, our results suggest that pushing students into work could slow down the educational progress of lower-performing students. Whether this shift is welfare enhancing depends on how long earnings increases last, how that compares to the cost of extra school years, and whether any of the marginal students are deterred from finishing high school (which likely has a large negative impact on future earnings).

## A.5 Robustness to Different Covariate Choices

The main text uses the post-double-selection LASSO (Belloni, Chernozhukov, and Hansen 2014a, 2014b; Belloni et al. 2012) to choose which covariates are included in each regression, as we pre-specified in our pre-analysis plan. For robustness, this section shows two different alternatives: including no covariates other than the cohort indicator needed for treatment to be conditionally random (i.e., controlling for randomization strata), and including all covariates that we feed into the post-double-selection process.

For employment outcomes, the covariates we feed into the lasso include indicators for: being male; being employed in each of the 2nd through 6th years prior to randomization; the earnings quartile of the pre-randomization year earnings; never being employed pre-SYEP; self-reporting being in high school, college, or being a high school graduate; being 15–16, 17–18, 19–20, or 21 and older; being part of the Ladders for Leaders program; being Hispanic, Asian, White, Other, or having missing race/ethnicity; not being matched to the education data; and being in the expected in high school sample.[2a]For the education outcomes, covariates we feed into the lasso include indicators for: being in grade 8 or under, grade 10, grade 11, or grade 12; being in deciles 1 through 9 of prior year GPA or missing GPA; being in quartiles 2 through 4 of the share of enrolled days attended; being male; being employed in each of the 2nd through 6th years prior to randomization; the earnings quartile of the pre-randomization year earnings; never being employed pre-SYEP; self-reporting being in high school, college, or being a high school graduate; being 15–16, 17–18, 19–20, or 21 and older; being part of the Ladders for Leaders program; and being Hispanic, Asian, White, Other, or having missing race/ethnicity.

Tables A.22 and A.23 show alternative results for labor market and education effects, respectively, controlling either for no covariates, other than the randomization stratum indicator needed for conditional independence, or all covariates. These tables lead to the same conclusions as the main tables. Because of the imbalance in several education baseline covariates discussed in section A.4.1, the point estimates on GPA and graduation measures become somewhat larger and more significant in specifications without covariate controls.

## A.6 Comparing Our Main Sample and Everyone on a Survey

The main text focuses on the sample of youth who were on a survey that a supervisor started, a group that we pre-specified as being of special interest in our pre-analysis plan. This excludes 25,813 young people who were only on surveys that no one started. Since none of these individuals could possibly have been treated if assigned to treatment, everyone in this group is effectively a never-taker. Since we are able to observe this fact for both treatment and control youth on these surveys, we exclude them from our main analysis to help with power.

This section provides some additional information on who is excluded from the sample and the implications for our analysis. Table A.24 compares our main control group to everyone who was on an unopened survey (treatment and control) on baseline characteristics. Given that assignment to supervisors was not random, it is not surprising that young people whose supervisors did not start the survey are observably different than those in our main sample. Table A.24 shows that our main sample is younger, less Black and less White (more Hispanic and Asian), more likely to still be in high school, and generally less engaged in the labor force pre-randomization than those on unopened surveys.

Table A.25 shows the same comparison but for outcome measures rather than baseline

2a. Ladders for Leaders is a special application-based program within the broader SYEP.

characteristics (which is why we only use the control group for those on a started survey). The table indicates that our control group continues to be less involved in the labor market than those on unopened surveys during the outcome period, but more engaged and successful in school. There is, however, no significant difference in job application behavior, consistent with the argument in the main text that differences in employment status do not affect the decision of whether to apply to our job.

Given the observable differences between our main sample and those on unopened surveys, our estimates are most externally valid for the group that would look most like those in our main sample: young people whose supervisors fill out the surveys when asked, without any requirement to do so. It is possible that forcing supervisors to fill out surveys for their employees could generate somewhat different effects, given that the population of youth affected would be observably different.

Table A.26 shows the main employment and earnings results for the full sample of everyone on a survey, rather than our main sample of everyone on a started survey. As we would expect from the inclusion of almost 26,000 additional never-takers, the estimates are somewhat less precise than our main results. But the patterns are quite similar and still statistically significant at the 0.1 level: an increase in year 1 employment that fades out over time, and an increase in earnings that grows in both levels and proportions over time to an additional $1,470 (5.3 percent) in cumulative earnings.

## A.7 Additional Figures and Tables

Figure A.1: Example Supervisor Survey Invitation Email

Dear Judd Kessler,

Thank you for your participation in the 2017 Summer Youth Employment Program (SYEP), run by the New York City Department of Youth and Community Development.

For the second year, we are running a "letter of recommendation" program. As part of this program, **we are asking you to complete a very short survey** about some of the youth who worked for you this summer (the survey should take about 1 minute per selected youth).

Positive responses will be turned into letters of recommendation for the youth. We expect these letters to help youth capitalize on their experience working for you this summer.

To join employers like you in participating, please click on this personalized link by **a week from tomorrow, Friday, October 20th**: Take the survey.

If you have any questions about the program, please see a further description on our website here.

If you have additional questions, you can contact our academic partners: Judd B. Kessler (judd.kessler@wharton.upenn.edu) at the University of Pennsylvania and Sara Heller (hellersa@sas.upenn.edu).

Sincerely,

SYEP Team

Follow the link to opt out of future emails:
Click here to unsubscribe

Figure A.2: Screen Shots from Beginning of Supervisor Survey

Judd Kessler, thank you for participating in the 2017 Summer Youth Employment Program, run by the New York City Department of Youth and Community Development.

For the second year, we are conducting a "letter of recommendation" program. We are asking you, and employers like you, to answer a very short survey about some of the youth who worked for you this summer. If you rate a youth positively, your responses will be turned into sentences and put into a recommendation letter from you on DYCD letterhead. The youth can then show this letter to future educators and potential employers. (If you are interested, see a sample letter here.)

So that we can ask you about the correct youth, please confirm the following information:

My name is Judd Kessler.

- ◯ Yes
- ◯ No

I supervised or worked with summer youth employees at the University of Pennsylvania.

- ◯ Yes
- ◯ No

››

The following youth from University of Pennsylvania have been randomly selected to participate in the program. Please select which youth you supervised or worked with this summer. If you did not supervise any youth this summer, please leave these boxes unchecked.

- ☐ Sara Heller
- ☐ Andre Padilla
- ☐ William Schmidt
- ☐ Fernando Willis

››

You have indicated that you supervised the following youth who have been randomly selected to participate in this program.

Sara Heller
Andre Padilla
William Schmidt
Fernando Willis

For each youth, we will ask you for an overall rating and give you the option to create a letter of recommendation for that youth.

If you choose to create a letter for a particular youth, you will also have the option to include your contact information so that their potential future employers can reach you as a reference.

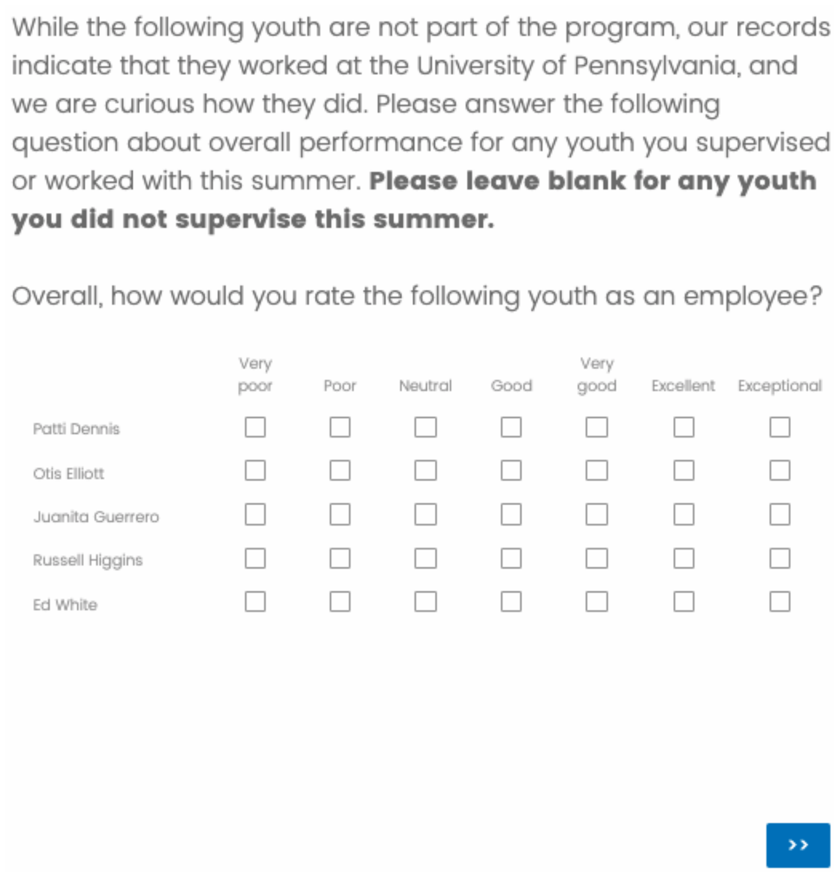
To be a reference for one or more of the youth listed above, please provide a phone number and/or email address so that potential future employers know how to reach you. (We will only ask for your information once, but we'll ask you if you want to be a reference for each youth separately. So even if you provide contact information here, it will only be included in letters for the youth you select.)

Phone number

Email address

››

Figure A.3: Screenshot of Control Youth Rating on Supervisor Survey



While the following youth are not part of the program, our records indicate that they worked at the University of Pennsylvania, and we are curious how they did. Please answer the following question about overall performance for any youth you supervised or worked with this summer. **Please leave blank for any youth you did not supervise this summer.**

Overall, how would you rate the following youth as an employee?

| | Very poor | Poor | Neutral | Good | Very good | Excellent | Exceptional |
|---|---|---|---|---|---|---|---|
| Patti Dennis | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Otis Elliott | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Juanita Guerrero | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Russell Higgins | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Ed White | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

››

Figure A.4: Example Cover Letter to the Letter of Recommendation

Department of
Youth & Community
Development

November 1, 2017

Sara Heller
123 Fake Street
New York, NY 10003

Dear Sara,

This past summer you participated in a New York City summer program. This letter contains five copies of a letter of recommendation your supervisor wrote for you. [You should also have received a link to an electronic copy at [Student Email], in case you want to have an electronic version or print out more of copies of the letter.]

This year, some participants were included in a "letter of recommendation" program. You were included in this program, and your employer gave us feedback that could help you get a job or show your teachers your strengths. We hope you will show your letter of recommendation to your teachers, your guidance counselor, and potential employers (for example, by including it in job applications).

If you have any questions about the program, please see a description on our website here: https://www1.nyc.gov/assets/dycd/downloads/pdf/FAQs_Pilot_2017.pdf

If you have additional questions, you can contact our academic partners: Judd B. Kessler (judd.kessler@wharton.upenn.edu) at the University of Pennsylvania, and Sara Heller (hellersa@sas.upenn.edu).

Sincerely,

DYCD Team

The New York City Department of Youth and Community Development (DYCD) invests in a network of community-based organizations and programs to alleviate the effects of poverty and to provide opportunities for New Yorkers and communities to flourish.

Empowering Individuals • Strengthening Families • Investing in Communities

Notes: This cover letter accompanied five copies of the recommendation sent to youth. The text in brackets appeared when we had an email address on file for the youth.

Figure A.5: Example Job Advertisement Email

Penn
University of Pennsylvania

## Youth Job Opening!

Dear **Sara**:

A professor at the University of Pennsylvania is looking for former NYC summer job program participants, like you, to **apply for a short-term and flexible job.**

This is an opportunity to **earn money and gain work experience** while helping improve future youth employment programs.

Those hired will **not** need to be on site at the University of Pennsylvania. All tasks and duties necessary to the position will be completed remotely (on an Internet capable computer or by mail).

**Qualifications:**

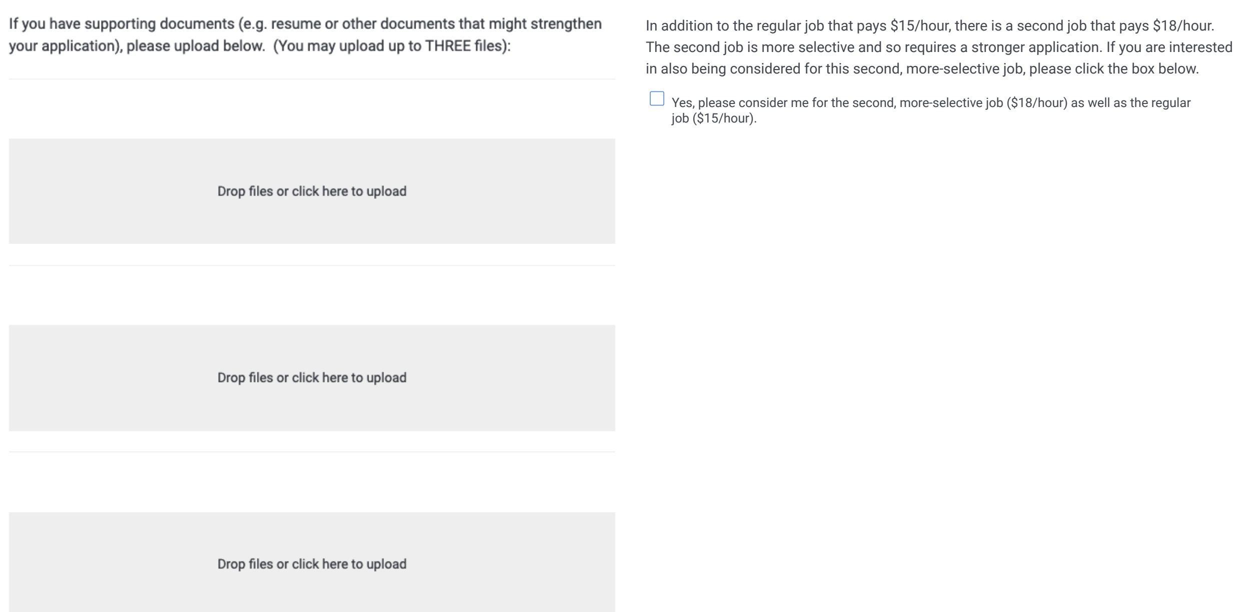
- Responsible
- Self-motivated
- Enthusiastic approach
- Some work experience preferred

**Compensation for the job is $15/hour.**

If you are **Sara Heller**, click this link to apply **(application due by March 30th)**:
Click here for your personal job application

**All others** who are interested can click this link for more information and to apply:
General job application

Figure A.6: Job Application Prompts to Upload Supporting Documents and to be Considered for More Selective Job



If you have supporting documents (e.g. resume or other documents that might strengthen your application), please upload below. (You may upload up to THREE files):

Drop files or click here to upload

Drop files or click here to upload

Drop files or click here to upload

In addition to the regular job that pays $15/hour, there is a second job that pays $18/hour. The second job is more selective and so requires a stronger application. If you are interested in also being considered for this second, more-selective job, please click the box below.

☐ Yes, please consider me for the second, more-selective job ($18/hour) as well as the regular job ($15/hour).

Figure A.7: Available 4th- to 6th-Year Graduation Data Relative to Randomization, by Grade and Study Cohort

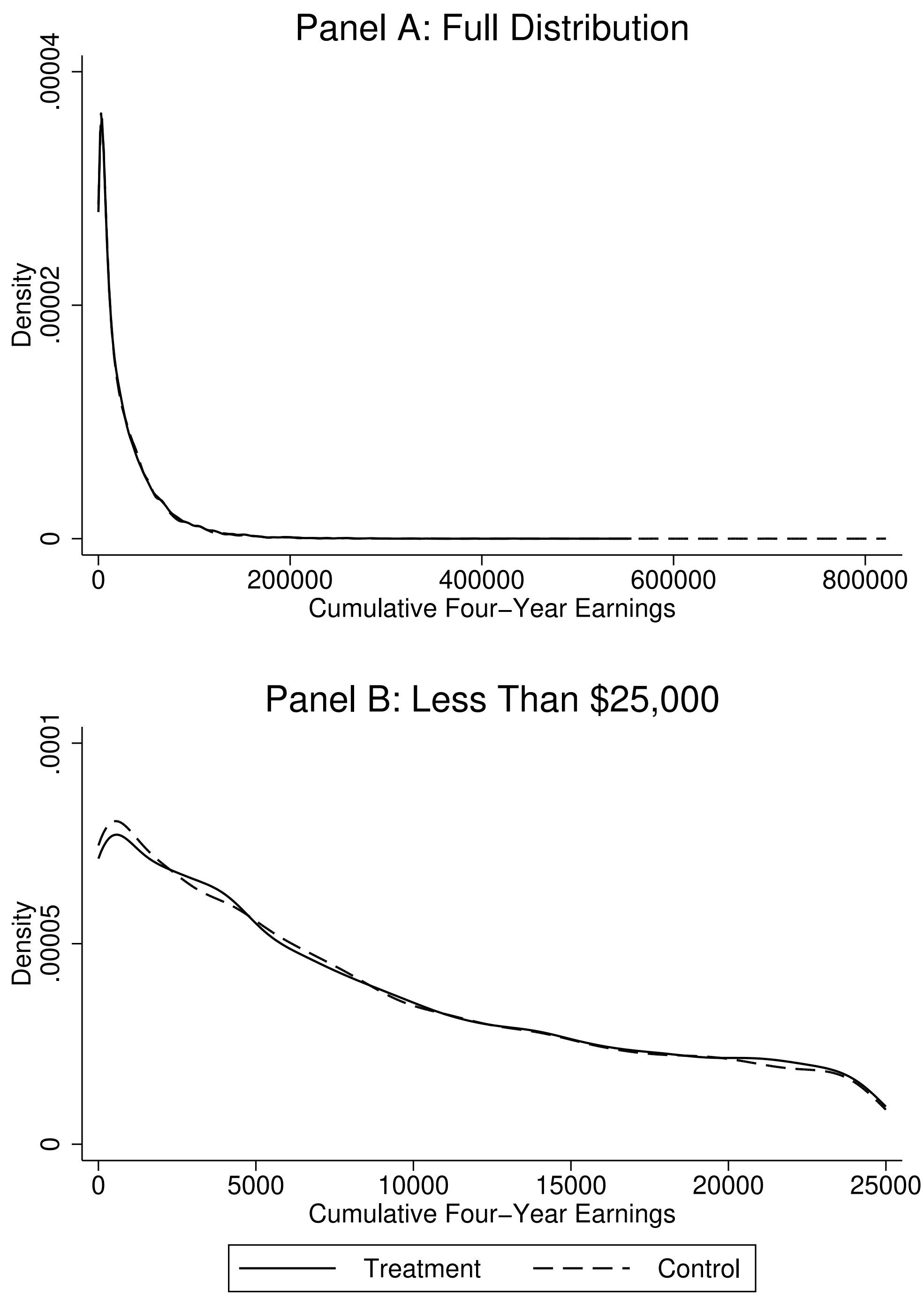
| Pre-Randomization Grade | 8th | 9th | 10th | 11th | 12th |
|---|---|---|---|---|---|
| Year Relative to Randomization | 2016 Study Cohort | | | | |
| | N = 268 | 994 | 1,313 | 1,459 | 149 |
| -1 (graduated by 8/2016) | | | | | 4th |
| 1 (by 8/2017) | | | | 4th | 5th |
| 2 (by 8/2018) | | | 4th | 5th | 6th |
| 3 (by 8/2019) | | 4th | 5th | 6th | |
| 4 (by 8/2020) | 4th | 5th | 6th | | |
| 5 (by 8/2021) | 5th | 6th | | | |
| | 2017 Study Cohort | | | | |
| | N = 1,177 | 3,543 | 4,984 | 5,249 | 578 |
| -1 (by 8/2017) | | | | | 4th |
| 1 (by 8/2018) | | | | 4th | 5th |
| 2 (by 8/2019) | | | 4th | 5th | 6th |
| 3 (by 8/2020) | | 4th | 5th | 6th | |
| 4 (by 8/2021) | 4th | 5th | 6th | | |

[black box] = Included in graduation measures

[gray box] = Not old enough to observe

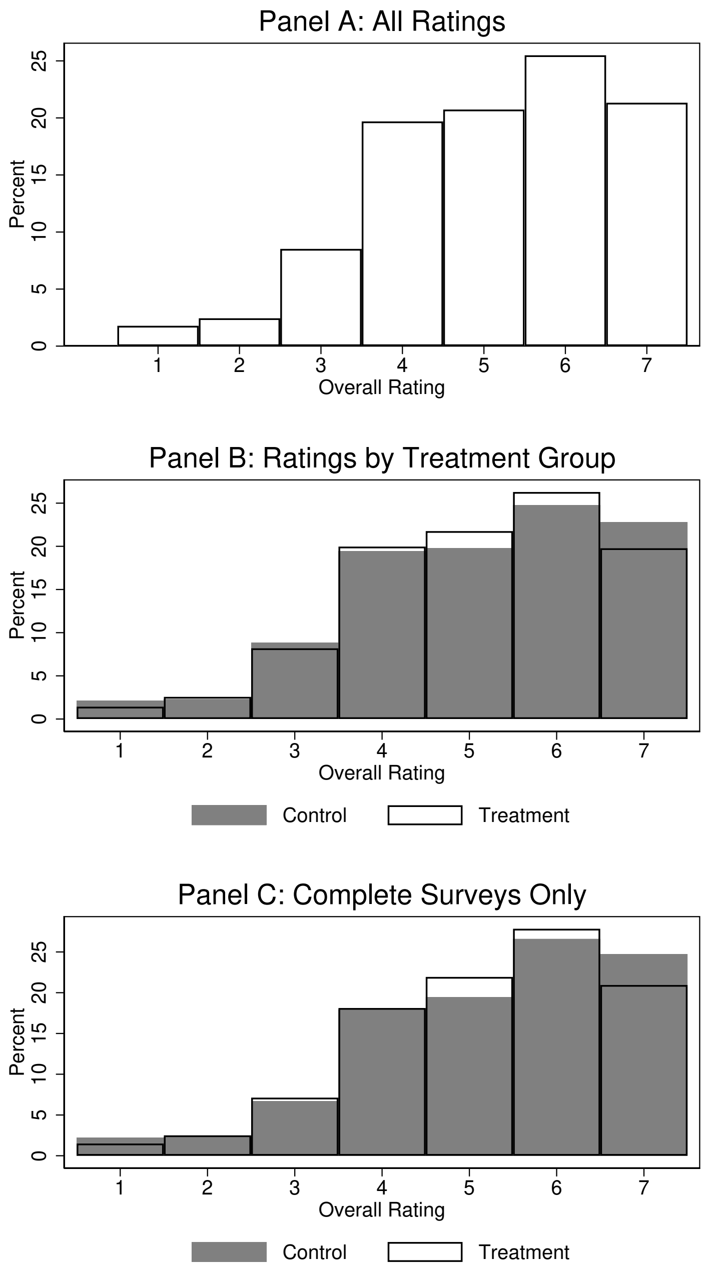
Notes: Figure shows when 4th-, 5th-, and 6th-year graduation outcomes are observed for students in each pre-randomization grade level by study cohort. Black boxes define our main "expected in high school" sample, for whom at least on-time graduation is observed. Gray boxes show the graduation outcomes that are not yet observed in our data. Only 12th graders who had not graduated prior to letter distribution are included in these samples, so they are all recorded as not having graduated by their 4th year graduation date in year -1.

Figure A.8: Cumulative Earnings Distribution




Notes: Figure shows the distribution of total earnings for treatment and control groups over 4 years, with one extreme outlier in one quarter (over $3 million) top-coded to equal the next highest quarterly amount in the data prior to summing over all quarters. Panel A shows the full distribution. Panel B zooms in on the lower end of the distribution to make treatment-control differences visible.

Figure A.9: Distribution of Supervisor Ratings




Notes: N = 29,877 for all surveys and 13,911 for completed surveys. Figure shows distribution of non-missing supervisor ratings for everyone (Panel A), separately by treatment group (Panel B), and by treatment group just for youth on fully-completed surveys (Panel C). Our main analysis maps categories 1–4 to "low" and categories 5–7 to "high."

Table A.1: Earnings Impacts across Different Skewness Adjustments

| Year | 1 | 2 | 3 | 4 | Cumulative |
|---|---|---|---|---|---|
| | Panel A: Winsorized at 99.5th Percentile | | | | |
| ITT | 57.96 | 104.37 | 128.83 | 214.72* | 544.52** |
| | (43.16) | (71.94) | (96.65) | (128.38) | (277.26) |
| CM | 3532 | 5925 | 7378 | 9927 | 26852 |
| Sent Letter (IV) | 149 | 267 | 330 | 546.06* | 1348.83** |
| | (106.66) | (177.76) | (238.80) | (317.24) | (685.56) |
| CCM | 3682 | 6132 | 7554 | 10239 | 27661 |
| | Panel B: Log(Earnings + 0.1) | | | | |
| ITT | 0.125*** | 0.073 | 0.042 | 0.016 | 0.048 |
| | (0.042) | (0.044) | (0.048) | (0.050) | (0.031) |
| CM | 4.92 | 5.44 | 4.86 | 5.36 | 8.68 |
| Sent Letter (IV) | 0.309*** | 0.18 | 0.109 | 0.04 | 0.12 |
| | (0.104) | (0.110) | (0.119) | (0.124) | (0.077) |
| CCM | 4.94 | 5.56 | 5.01 | 5.55 | 8.76 |
| | Panel C: Log(Earnings + 1) | | | | |
| ITT | 0.095*** | 0.059* | 0.035 | 0.013 | 0.042* |
| | (0.033) | (0.035) | (0.038) | (0.040) | (0.026) |
| CM | 5.61 | 6.08 | 5.67 | 6.09 | 8.86 |
| Sent Letter (IV) | 0.236*** | 0.145* | 0.091 | 0.033 | 0.105* |
| | (0.081) | (0.087) | (0.095) | (0.100) | (0.063) |
| CCM | 5.64 | 6.18 | 5.79 | 6.25 | 8.94 |
| | Panel D: Log(Earnings + 10) | | | | |
| ITT | 0.066*** | 0.045* | 0.028 | 0.012 | 0.035* |
| | (0.024) | (0.026) | (0.029) | (0.031) | (0.021) |
| CM | 6.30 | 6.73 | 6.48 | 6.83 | 9.04 |
| Sent Letter (IV) | 0.164*** | 0.111* | 0.069 | 0.028 | 0.088* |
| | (0.058) | (0.064) | (0.071) | (0.076) | (0.051) |
| CCM | 6.34 | 6.81 | 6.58 | 6.95 | 9.11 |
| | Panel E: Log(Earnings + 100) | | | | |
| ITT | 0.038** | 0.031* | 0.02 | 0.01 | 0.028* |
| | (0.015) | (0.017) | (0.020) | (0.021) | (0.016) |
| CM | 7.03 | 7.40 | 7.31 | 7.59 | 9.24 |
| Sent Letter (IV) | 0.095** | 0.076* | 0.05 | 0.02 | 0.070* |
| | (0.037) | (0.043) | (0.048) | (0.053) | (0.039) |
| CCM | 7.07 | 7.47 | 7.38 | 7.68 | 9.30 |
| | Panel F: Asinh(Earnings) | | | | |
| ITT | 0.104*** | 0.063* | 0.037 | 0.015 | 0.044 |
| | (0.035) | (0.038) | (0.041) | (0.043) | (0.027) |
| CM | 6.09 | 6.58 | 6.12 | 6.56 | 9.50 |
| Sent Letter (IV) | 0.258*** | 0.155* | 0.097 | 0.036 | 0.109 |
| | (0.088) | (0.094) | (0.102) | (0.107) | (0.067) |
| CCM | 6.12 | 6.69 | 6.25 | 6.73 | 9.58 |

Notes: N = 43,409. Winsorization in Panel A recodes each quarter's highest earnings to the 99.5th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.2: Spell Length and Censoring

| | Total Spell Length | Spell Censored | Lasts at Least 2 Qtrs | Lasts at Least 3 Qtrs | Lasts at Least 4 Qtrs |
|---|---|---|---|---|---|
| | | | Spell 1 | | |
| ITT | 0.0331 | -0.0024 | 0.0019 | 0.0046 | 0.0004 |
| | (0.0365) | (0.0025) | (0.0045) | (0.0047) | (0.0045) |
| CM | 3.72 | 0.07 | 0.61 | 0.45 | 0.35 |
| IV | 0.0825 | -0.0060 | 0.0047 | 0.0114 | 0.0013 |
| | (0.0897) | (0.0061) | (0.0111) | (0.0116) | (0.0111) |
| CCM | 3.93 | 0.07 | 0.63 | 0.49 | 0.38 |
| N | 40088 | 40088 | 39537 | 39159 | 38914 |
| | | | Spell 2 | | |
| ITT | 0.0412 | -0.0046 | 0.0021 | 0.0078 | 0.005 |
| | (0.0279) | (0.0037) | (0.0054) | (0.0053) | (0.0048) |
| CM | 2.71 | 0.15 | 0.57 | 0.36 | 0.25 |
| IV | 0.10 | -0.0111 | 0.0053 | 0.0193 | 0.0124 |
| | (0.0678) | (0.0091) | (0.0131) | (0.0130) | (0.0117) |
| CCM | 2.78 | 0.16 | 0.59 | 0.38 | 0.26 |
| N | 34228 | 34228 | 32737 | 31769 | 31126 |
| | | | Spell 3 | | |
| ITT | 0.0149 | -0.0070 | 0.0009 | 0.0023 | -0.0024 |
| | (0.0274) | (0.0051) | (0.0063) | (0.0064) | (0.0059) |
| CM | 2.51 | 0.24 | 0.60 | 0.37 | 0.25 |
| IV | 0.0370 | -0.0172 | 0.0023 | 0.0051 | -0.0062 |
| | (0.0660) | (0.0123) | (0.0152) | (0.0153) | (0.0141) |
| CCM | 2.56 | 0.25 | 0.61 | 0.38 | 0.26 |
| N | 26099 | 26099 | 23849 | 22545 | 21556 |

Notes: Total Spell Length conditions on youth having a spell. Censored is an indicator for working in a spell in the last quarter observed. Indicators for at least X quarters are conditional on observing at least X quarters in the data. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.3: Labor Market Effects for DYCD and Non-DYCD Employers

| | DYCD | | | | | Non-DYCD Employers | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | 1 | 2 | 3 | 4 | Cumulative | 1 | 2 | 3 | 4 | Cumulative |
| | | | | | Panel A: Employment | | | | | |
| ITT | 0.0049 | 0.0034 | -0.0022 | 0.0000 | 0.0003 | 0.0088** | 0.0009 | 0.0023 | 0.0014 | 0.0037 |
| | (0.0046) | (0.0041) | (0.0020) | (0.0024) | (0.0046) | (0.0040) | (0.0043) | (0.0044) | (0.0045) | (0.0035) |
| CM | 0.4158 | 0.2619 | 0.052 | 0.0698 | 0.5227 | 0.4254 | 0.5654 | 0.6252 | 0.6494 | 0.8269 |
| Sent Letter (IV) | 0.0124 | 0.0081 | -0.0054 | 0.0000 | 0.001 | 0.0221** | 0.0023 | 0.0063 | 0.0041 | 0.0091 |
| | (0.0114) | (0.0101) | (0.0050) | (0.0059) | (0.0114) | (0.0100) | (0.0106) | (0.0108) | (0.0110) | (0.0085) |
| CCM | 0.419 | 0.253 | 0.056 | 0.070 | 0.531 | 0.429 | 0.588 | 0.638 | 0.664 | 0.836 |
| | | | | | Panel B: Earnings, Winsorized at 99th Percentile | | | | | |
| ITT | 1.22 | 3.35 | -4.48 | 2.63 | 9.50 | 58.46 | 99.84 | 132.97 | 212.58* | 506.15* |
| | (10.68) | (9.94) | (5.16) | (4.87) | (34.59) | (43.54) | (72.60) | (96.82) | (128.62) | (274.77) |
| CM | 810 | 572 | 117 | 131 | 2527 | 2724 | 5353 | 7261 | 9796 | 25134 |
| Sent Letter (IV) | 2.28 | 7.94 | -11.03 | 6.48 | 26.88 | 144.72 | 253.76 | 341.78 | 540.39* | 1290.77* |
| | (26.42) | (24.56) | (12.78) | (12.06) | (85.42) | (107.56) | (179.44) | (239.22) | (317.84) | (678.88) |
| CCM | 870 | 574 | 128 | 128 | 2592 | 2816 | 5563 | 7426 | 10111 | 25905 |

N = 43,409. DYCD shows employment and earnings at employers with the FEIN of the agency that runs the SYEP. Non-DYCD shows all other employment. Winsorization in Panel B recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.4: Labor Market Effects by Industry Cluster

| | SYEP-Related Industries (Cluster 1) | | | | | Other Industries (Cluster 2) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | 1 | 2 | 3 | 4 | Cumulative | 1 | 2 | 3 | 4 | Cumulative |
| | Panel A: Employment | | | | | | | | | |
| ITT | 0.0047 | 0.0063 | 0.0001 | -0.0013 | 0.0017 | 0.0071* | 0.0013 | 0.0029 | 0.0000 | -0.0007 |
| | (0.0047) | (0.0047) | (0.0042) | (0.0043) | (0.0042) | (0.0040) | (0.0044) | (0.0046) | (0.0047) | (0.0042) |
| CM | 0.5243 | 0.4407 | 0.2832 | 0.3092 | 0.7252 | 0.3104 | 0.4246 | 0.468 | 0.4879 | 0.7043 |
| Sent Letter (IV) | 0.0117 | 0.0151 | 0.0003 | -0.0029 | 0.004 | 0.0180* | 0.0036 | 0.0076 | 0.0000 | -0.0013 |
| | (0.0116) | (0.0116) | (0.0104) | (0.0108) | (0.0104) | (0.0099) | (0.0110) | (0.0114) | (0.0116) | (0.0104) |
| CCM | 0.537 | 0.443 | 0.301 | 0.33 | 0.743 | 0.307 | 0.435 | 0.468 | 0.493 | 0.713 |
| | Panel B: Earnings, Winsorized at 99th Percentile | | | | | | | | | |
| ITT | 17.06 | -2.41 | -54.79 | 77.61 | 36.25 | 35.34 | 104.35 | 193.36** | 158.52 | 487.14** |
| | (29.72) | (48.31) | (68.03) | (91.12) | (193.74) | (39.11) | (64.54) | (83.22) | (109.27) | (236.85) |
| CM | 1645 | 2242 | 2580 | 3576 | 10043 | 1853 | 3614 | 4689 | 6224 | 16380 |
| Sent Letter (IV) | 46.17 | -5.44 | -137.75 | 190.07 | 102.83 | 88.04 | 264.07* | 487.93** | 403.5 | 1233.96** |
| | (73.43) | (119.46) | (168.22) | (225.21) | (478.99) | (96.71) | (159.56) | (205.75) | (270.10) | (585.23) |
| CCM | 1802 | 2441 | 2849 | 3804 | 10887 | 1855 | 3603 | 4542 | 6229 | 16240 |

N = 43,409. Industry definition follows the cluster definitions in Gelber, Isen, and Kessler 2016. SYEP-related include employment in industries that are over-represented among summer jobs in the program. Other industries are those under-represented in summer jobs. Winsorization in Panel B recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.5: Balance for All Rated Youth by Rating Group

| | Control Low | Treatment Low | Test of Difference | Control High | Treatment High | Test of Difference |
|---|---|---|---|---|---|---|
| N | 5062 | 4632 | | 10425 | 9768 | |
| Age | 17.14 | 17.06 | 0.084 | 17.25 | 17.25 | 1.000 |
| Male | 0.449 | 0.448 | 0.935 | 0.414 | 0.417 | 0.753 |
| Black | 0.492 | 0.500 | 0.419 | 0.382 | 0.371 | 0.118 |
| Hispanic | 0.292 | 0.294 | 0.836 | 0.284 | 0.287 | 0.678 |
| Asian | 0.099 | 0.091 | 0.224 | 0.147 | 0.159 | 0.015 |
| White | 0.069 | 0.070 | 0.875 | 0.140 | 0.137 | 0.566 |
| Other Race | 0.049 | 0.045 | 0.392 | 0.047 | 0.045 | 0.598 |
| In High School | 0.782 | 0.787 | 0.549 | 0.739 | 0.734 | 0.371 |
| HS Graduate | 0.046 | 0.043 | 0.469 | 0.040 | 0.040 | 0.768 |
| In College | 0.133 | 0.132 | 0.950 | 0.204 | 0.208 | 0.390 |
| Not in UI Data | 0.006 | 0.006 | 0.862 | 0.003 | 0.003 | 0.440 |
| Never Employed Pre-SYEP | 0.461 | 0.489 | 0.007 | 0.438 | 0.437 | 0.861 |
| Ever Worked, Year -4 | 0.144 | 0.129 | 0.026 | 0.159 | 0.159 | 0.963 |
| Earnings, Year -4 | 258 | 254 | 0.875 | 332 | 341 | 0.715 |
| Ever Worked, Year -3 | 0.254 | 0.236 | 0.037 | 0.274 | 0.282 | 0.178 |
| Earnings, Year -3 | 492 | 463 | 0.375 | 613 | 630 | 0.591 |
| Ever Worked, Year -2 | 0.424 | 0.403 | 0.035 | 0.450 | 0.454 | 0.524 |
| Earnings, Year -2 | 974 | 862 | 0.017 | 1104 | 1124 | 0.612 |
| Ever Worked, Year -1 | 0.964 | 0.978 | 0.000 | 0.989 | 0.993 | 0.012 |
| Earnings, Year -1 | 2169 | 2101 | 0.209 | 2478 | 2520 | 0.334 |
| No Education Match | 0.094 | 0.089 | 0.399 | 0.131 | 0.130 | 0.846 |
| In HS Sample | 0.488 | 0.494 | 0.542 | 0.440 | 0.441 | 0.826 |
| Joint F-test | F(24, 9587) = 1.382, p=.101 | | | F(24, 19643) = .711, p=.846 | | |

Notes: Sample includes all youth with employer rating (N = 29,887, 256 youth missing race/ethnicity). Low includes rating categories 1–4; High includes rating categories 5–7. Test of difference reports the p-value from a regression of each characteristic on a treatment indicator within that rating group, controlling for a cohort indicator and using standard errors clustered on individual.

Table A.6: Balance by Rating Group, Fully Completed Surveys

| | Control Low | Treatment Low | Test of Difference | Control High | Treatment High | Test of Difference |
|---|---|---|---|---|---|---|
| N | 2209 | 2092 | | 4833 | 4777 | |
| Age | 17.09 | 17.09 | 0.919 | 17.26 | 17.23 | 0.453 |
| Male | 0.440 | 0.439 | 0.937 | 0.400 | 0.409 | 0.352 |
| Black | 0.505 | 0.535 | 0.053 | 0.388 | 0.381 | 0.481 |
| Hispanic | 0.277 | 0.258 | 0.178 | 0.286 | 0.292 | 0.491 |
| Asian | 0.117 | 0.111 | 0.573 | 0.165 | 0.178 | 0.090 |
| White | 0.051 | 0.047 | 0.465 | 0.114 | 0.105 | 0.145 |
| Other Race | 0.050 | 0.049 | 0.874 | 0.047 | 0.044 | 0.461 |
| In High School | 0.785 | 0.783 | 0.846 | 0.735 | 0.736 | 0.941 |
| HS Graduate | 0.042 | 0.041 | 0.808 | 0.037 | 0.031 | 0.141 |
| In College | 0.137 | 0.139 | 0.890 | 0.211 | 0.216 | 0.550 |
| Not in UI Data | 0.005 | 0.007 | 0.595 | 0.003 | 0.004 | 0.352 |
| Never Employed Pre-SYEP | 0.481 | 0.482 | 0.942 | 0.450 | 0.463 | 0.222 |
| Ever Worked, Year -4 | 0.134 | 0.125 | 0.390 | 0.150 | 0.149 | 0.963 |
| Earnings, Year -4 | 238 | 232 | 0.887 | 326 | 308 | 0.578 |
| Ever Worked, Year -3 | 0.242 | 0.234 | 0.538 | 0.259 | 0.266 | 0.394 |
| Earnings, Year -3 | 439 | 447 | 0.864 | 596 | 582 | 0.749 |
| Ever Worked, Year -2 | 0.407 | 0.394 | 0.415 | 0.435 | 0.426 | 0.388 |
| Earnings, Year -2 | 909 | 803 | 0.110 | 1041 | 1035 | 0.908 |
| Ever Worked, Year -1 | 0.970 | 0.979 | 0.065 | 0.989 | 0.991 | 0.333 |
| Earnings, Year -1 | 2075 | 2007 | 0.361 | 2427 | 2358 | 0.244 |
| No Education Match | 0.083 | 0.078 | 0.552 | 0.111 | 0.114 | 0.621 |
| In HS Sample | 0.498 | 0.493 | 0.736 | 0.460 | 0.455 | 0.660 |
| Joint F-test | F(24, 4264) = .889, p=.618 | | | F(24, 9471) = .862, p=.658 | | |

Notes: Sample includes all youth on a fully completed survey (N=13,911, 167 youth missing race/ethnicity). Low includes rating categories 1–4; High includes rating categories 5–7. Test of difference reports the p-value from a regression of each characteristic on a treatment indicator within that rating group, controlling for a cohort indicator and using standard errors clustered on individual.

Table A.7: Employment and Earnings Effects by Rating, On Any Survey

| | | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | | | Employment | | |
| ITT, Low Ratings | | 0.0135 | 0.0001 | -0.0101 | 0.0099 | 0.0097* |
| | | (0.0089) | (0.0089) | (0.0092) | (0.0094) | (0.0054) |
| ITT, High Ratings | | 0.0122** | 0.0100* | 0.0068 | 0.0008 | 0.0003 |
| | | (0.0059) | (0.0060) | (0.0063) | (0.0063) | (0.0035) |
| P-value, test of diff. | | 0.906 | 0.358 | 0.131 | 0.418 | 0.145 |
| CM, Low | | 0.679 | 0.711 | 0.653 | 0.664 | 0.915 |
| CM, High | | 0.724 | 0.732 | 0.659 | 0.694 | 0.931 |
| | First Stage | | | | | |
| IV, Low Ratings | 0.3067*** | 0.0435 | 0.0002 | -0.0331 | 0.0324 | 0.0316* |
| | (0.0068) | (0.0290) | (0.0290) | (0.0301) | (0.0307) | (0.0177) |
| IV, High Ratings | 0.7529*** | 0.0161** | 0.0132* | 0.0089 | 0.001 | 0.0003 |
| | (0.0043) | (0.0079) | (0.0079) | (0.0084) | (0.0084) | (0.0047) |
| P-value, test of diff. | 0.000 | 0.362 | 0.665 | 0.179 | 0.323 | 0.088 |
| CCM, Low | | 0.65 | 0.726 | 0.685 | 0.660 | 0.895 |
| CCM, High | | 0.719 | 0.733 | 0.664 | 0.700 | 0.931 |
| | | | | Earnings, Winsorized at 99th Percentile | | |
| ITT, Low Ratings | | 0.41 | 2.43 | -25.25 | -126.73 | -142.21 |
| | | (84.53) | (140.43) | (181.03) | (237.26) | (518.51) |
| ITT, High Ratings | | 116.09* | 175.34 | 71.40 | 293.61 | 642.02 |
| | | (65.73) | (109.88) | (148.49) | (198.99) | (428.95) |
| P-value, test of diff. | | 0.280 | 0.332 | 0.68 | 0.175 | 0.244 |
| CM, Low | | 3202 | 5418 | 6598 | 8629 | 23884 |
| CM, High | | 3778 | 6292 | 7942 | 10752 | 28874 |
| | First Stage | | | | | |
| IV, Low Ratings | 0.3067*** | -17.41 | 12.81 | -82.44 | -407.86 | -464.03 |
| | (0.0068) | (274.99) | (457.96) | (590.58) | (774.24) | (1691.22) |
| IV, High Ratings | 0.7529*** | 151.68* | 234.65 | 94.80 | 387.30 | 852.53 |
| | (0.0043) | (87.20) | (145.95) | (197.18) | (264.17) | (569.70) |
| P-value, test of diff. | 0 | 0.558 | 0.644 | 0.776 | 0.331 | 0.461 |
| CCM, Low | | 3204 | 5562 | 6676 | 8836 | 24273 |
| CCM, High | | 3804 | 6324 | 8039 | 10853 | 29161 |

Notes: N = 29,887. Sample includes all youth with ratings, regardless of whether supervisor completed all the ratings on the survey. Low includes rating categories 1–4; High includes rating categories 5–7. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.8: Employment and Earnings Effects by Expected in HS Sample

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | Employment | | | | |
| ITT, Expected | | 0.0145** | 0.0126* | 0.0066 | 0.0058 | 0.0048 |
| in High School | | (0.0065) | (0.0064) | (0.0066) | (0.0066) | (0.0039) |
| ITT, Not Expected | | 0.0113** | 0.0002 | -0.0002 | -0.0028 | 0.0012 |
| in High School | | (0.0052) | (0.0053) | (0.0057) | (0.0059) | (0.0032) |
| P-value, test of diff. | | 0.698 | 0.139 | 0.441 | 0.33 | 0.472 |
| CM, Exp. in HS | | 0.635 | 0.677 | 0.611 | 0.673 | 0.913 |
| CM, Not Exp. in HS | | 0.755 | 0.756 | 0.683 | 0.689 | 0.930 |
| | First Stage | | | | | |
| IV, Expected | 0.4138*** | 0.0351** | 0.0304* | 0.0169 | 0.0141 | 0.0119 |
| in High School | (0.0049) | (0.0158) | (0.0155) | (0.0161) | (0.0159) | (0.0094) |
| IV, Not Expected | 0.3966*** | 0.0285** | 0.0005 | -0.0004 | -0.0069 | 0.0029 |
| in High School | (0.0044) | (0.0132) | (0.0135) | (0.0143) | (0.0148) | (0.0082) |
| P-value, test of diff. | 0.010 | 0.745 | 0.146 | 0.42 | 0.333 | 0.471 |
| CCM, Exp. in HS | | 0.643 | 0.68 | 0.611 | 0.678 | 0.917 |
| CCM, Not Exp. in HS | | 0.743 | 0.770 | 0.706 | 0.713 | 0.929 |
| | | Earnings, Winsorized at 99th Percentile | | | | |
| ITT, Expected | | 15.95 | 76.62 | 160.69 | 236.27* | 543.09* |
| in High School | | (40.95) | (76.23) | (104.62) | (142.76) | (291.48) |
| ITT, Not Expected | | 95.15 | 125.07 | 102.31 | 193.45 | 545.72 |
| in High School | | (71.46) | (115.40) | (154.25) | (203.27) | (446.95) |
| P-value, test of diff. | | 0.336 | 0.726 | 0.754 | 0.863 | 0.996 |
| CM, Exp. in HS | | 2097 | 3889 | 4952 | 7124 | 18077 |
| CM, Not Exp. in HS | | 4727 | 7620 | 9399 | 12261 | 34158 |
| | First Stage | | | | | |
| IV, Expected | 0.4138*** | 47.67 | 203.25 | 413.50 | 607.46* | 1323.39* |
| in High School | (0.0049) | (99.06) | (184.31) | (252.57) | (346.02) | (705.21) |
| IV, Not Expected | 0.3966*** | 236.58 | 315.61 | 258.62 | 486.65 | 1370.80 |
| in High School | (0.0044) | (179.71) | (290.65) | (388.61) | (511.35) | (1125.81) |
| P-value, test of diff. | 0.010 | 0.358 | 0.744 | 0.739 | 0.845 | 0.972 |
| CCM, Exp. in HS | | 2275 | 4063 | 4954 | 6907 | 18171 |
| CCM, Not Exp. in HS | | 4895 | 7922 | 9796 | 13119 | 35842 |

Notes: N = 43,409 (19,714 expected in high school data, 23,695 out of school or not expected in later education data). Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.9: Employment and Earnings Effects by Age

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | Employment | | | | |
| ITT, Under 18 | | 0.0137** | 0.0087 | 0.0056 | 0.0054 | 0.0045 |
| | | (0.0055) | (0.0054) | (0.0057) | (0.0055) | (0.0033) |
| ITT, 18 and Over | | 0.0110* | 0.0002 | -0.0023 | -0.0066 | -0.0002 |
| | | (0.0060) | (0.0063) | (0.0068) | (0.0071) | (0.0039) |
| P-value, test of diff. | | 0.734 | 0.306 | 0.369 | 0.185 | 0.347 |
| CM, Under 18 | | 0.645 | 0.685 | 0.601 | 0.677 | 0.916 |
| CM, 18 and Over | | 0.798 | 0.780 | 0.735 | 0.691 | 0.934 |
| | First Stage | | | | | |
| IV, Under 18 | 0.4027*** | 0.0341** | 0.0216 | 0.0143 | 0.0133 | 0.0114 |
| | (0.0042) | (0.0136) | (0.0134) | (0.0140) | (0.0138) | (0.0081) |
| IV, 18 and Over | 0.4072*** | 0.0267* | 0.0004 | -0.0049 | -0.0163 | -0.0009 |
| | (0.0055) | (0.0148) | (0.0155) | (0.0166) | (0.0175) | (0.0095) |
| P-value, test of diff. | 0.508 | 0.712 | 0.302 | 0.378 | 0.184 | 0.325 |
| CCM, Under 18 | | 0.647 | 0.695 | 0.613 | 0.684 | 0.917 |
| CCM, 18 and Over | | 0.783 | 0.787 | 0.747 | 0.719 | 0.936 |
| | | Earnings, Winsorized at 99th Percentile | | | | |
| ITT, Under 18 | | 45.28 | 60.70 | 127.61 | 147.54 | 387.72 |
| | | (35.72) | (65.56) | (90.20) | (122.61) | (251.03) |
| ITT, 18 and Over | | 80.62 | 184.34 | 133.07 | 328.55 | 826.31 |
| | | (100.70) | (161.77) | (213.61) | (279.18) | (622.25) |
| P-value, test of diff. | | 0.741 | 0.478 | 0.981 | 0.553 | 0.513 |
| CM, Under 18 | | 2120 | 3962 | 4977 | 7326 | 18404 |
| CM, 18 and Over | | 5979 | 9329 | 11542 | 14438 | 41501 |
| | First Stage | | | | | |
| IV, Under 18 | 0.4027*** | 116.87 | 150.74 | 338.71 | 375.06 | 984.70 |
| | (0.0042) | (88.61) | (162.85) | (223.49) | (304.75) | (623.29) |
| IV, 18 and Over | 0.4072*** | 201.90 | 452.80 | 328.53 | 836.23 | 2003.98 |
| | (0.0055) | (247.08) | (397.48) | (524.70) | (684.86) | (1528.46) |
| P-value, test of diff. | 0.508 | 0.746 | 0.482 | 0.986 | 0.539 | 0.537 |
| CCM, Under 18 | | 2243 | 4180 | 5060 | 7311 | 18813 |
| CCM, 18 and Over | | 6135 | 9467 | 11788 | 15225 | 42687 |

Notes: N = 43,409 (27,500 under 18, 15,909 age 18 and up). Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.10: Employment and Earnings Effects by Pre-SYEP Work Experience Status

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | | | Employment | | |
| ITT, Never Worked | | 0.0068 | 0.0052 | 0.0005 | 0.0037 | 0.0028 |
| | | (0.0068) | (0.0067) | (0.0068) | (0.0067) | (0.0043) |
| ITT, Ever Worked | | 0.0176*** | 0.0064 | 0.005 | -0.0009 | 0.0033 |
| | | (0.0050) | (0.0051) | (0.0056) | (0.0058) | (0.0028) |
| P-value, test of diff. | | 0.201 | 0.888 | 0.601 | 0.602 | 0.922 |
| CM, Never Worked | | 0.588 | 0.634 | 0.557 | 0.646 | 0.892 |
| CM, Ever Worked | | 0.793 | 0.790 | 0.727 | 0.711 | 0.947 |
| | First Stage | | | | | |
| IV, Never Worked | 0.3950*** | 0.0173 | 0.0132 | 0.0022 | 0.0094 | 0.0067 |
| | (0.0049) | (0.0173) | (0.0170) | (0.0171) | (0.0170) | (0.0110) |
| IV, Ever Worked | 0.4121*** | 0.0428*** | 0.0155 | 0.0121 | -0.0024 | 0.0074 |
| | (0.0045) | (0.0121) | (0.0123) | (0.0136) | (0.0140) | (0.0068) |
| P-value, test of diff. | 0.010 | 0.226 | 0.912 | 0.649 | 0.591 | 0.953 |
| CCM, Never Worked | | 0.601 | 0.646 | 0.573 | 0.652 | 0.896 |
| CCM, Ever Worked | | 0.775 | 0.795 | 0.735 | 0.732 | 0.946 |
| | | | | Earnings, Winsorized at 99th Percentile | | |
| ITT, Never Worked | | 36.57 | 6.37 | 9.42 | 170.96 | 267.40 |
| | | (37.83) | (74.50) | (101.95) | (143.66) | (283.22) |
| ITT, Ever Worked | | 69.35 | 177.93 | 232.79 | 253.07 | 747.62* |
| | | (72.48) | (115.96) | (155.12) | (202.61) | (448.61) |
| P-value, test of diff. | | 0.689 | 0.213 | 0.228 | 0.741 | 0.365 |
| CM, Never Worked | | 1745 | 3461 | 4459 | 6869 | 16547 |
| CM, Ever Worked | | 4993 | 7941 | 9766 | 12429 | 35282 |
| | First Stage | | | | | |
| IV, Never Worked | 0.3950*** | 103.08 | 34.28 | 43.97 | 469.71 | 671.08 |
| | (0.0049) | (95.48) | (188.20) | (257.23) | (363.49) | (715.99) |
| IV, Ever Worked | 0.4121*** | 170.13 | 433.03 | 558.67 | 600.76 | 1822.65* |
| | (0.0045) | (175.90) | (281.41) | (376.46) | (491.28) | (1089.21) |
| P-value, test of diff. | 0.010 | 0.738 | 0.239 | 0.258 | 0.830 | 0.377 |
| CCM, Never Worked | | 1920 | 3848 | 4878 | 7018 | 17673 |
| CCM, Ever Worked | | 5123 | 8000 | 9721 | 12850 | 35810 |

Notes: N = 43,409 (23,718 with work experience prior to the SYEP summer, 19,691 youth who never worked prior to the SYEP summer). Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.11: Employment and Earnings Effects for Minority and White Youth

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | Employment | | | | |
| ITT, Minority | | 0.0134*** | 0.0065 | 0.005 | 0.0028 | 0.0045* |
| | | (0.0044) | (0.0044) | (0.0046) | (0.0047) | (0.0027) |
| ITT, White | | 0.0049 | -0.0009 | -0.0098 | -0.0074 | -0.0086 |
| | | (0.0114) | (0.0122) | (0.0128) | (0.0129) | (0.0076) |
| P-value, test of diff. | | 0.486 | 0.566 | 0.279 | 0.461 | 0.102 |
| CM, Minority | | 0.6929 | 0.7228 | 0.6606 | 0.6901 | 0.9228 |
| CM, White | | 0.7514 | 0.6941 | 0.5698 | 0.617 | 0.9169 |
| | First Stage | | | | | |
| IV, Minority | 0.4188*** | 0.0321*** | 0.0156 | 0.0125 | 0.0067 | 0.0103 |
| | (0.0036) | (0.0106) | (0.0105) | (0.0111) | (0.0112) | (0.0064) |
| IV, White | 0.2973*** | 0.0163 | -0.0045 | -0.0327 | -0.0249 | -0.0287 |
| | (0.0088) | (0.0385) | (0.0412) | (0.0433) | (0.0437) | (0.0255) |
| P-value, test of diff. | 0.000 | 0.691 | 0.636 | 0.313 | 0.485 | 0.138 |
| CCM, Minority | | 0.691 | 0.730 | 0.665 | 0.700 | 0.921 |
| CCM, White | | 0.752 | 0.712 | 0.621 | 0.646 | 0.951 |
| | | Earnings, Winsorized at 99th Percentile | | | | |
| ITT, Minority | | 75.47* | 138.63* | 190.64* | 322.79** | 773.64*** |
| | | (45.72) | (76.32) | (101.51) | (133.81) | (289.22) |
| ITT, White | | -55.08 | -129.91 | -292.07 | -351.56 | -839.26 |
| | | (131.93) | (213.57) | (296.57) | (402.44) | (874.60) |
| P-value, test of diff. | | 0.350 | 0.236 | 0.123 | 0.112 | 0.080 |
| CM, Minority | | 3500 | 5926 | 7339 | 9726 | 26560 |
| CM, White | | 3662 | 5636 | 7202 | 10354 | 27023 |
| | First Stage | | | | | |
| IV, Minority | 0.4188*** | 185.55* | 339.82* | 467.05* | 786.77** | 1846.53*** |
| | (0.0036) | (109.08) | (182.13) | (242.24) | (319.28) | (690.67) |
| IV, White | 0.2973*** | -192.20 | -454.73 | -977.60 | -1186.32 | -2865.06 |
| | (0.0088) | (444.91) | (720.07) | (1001.75) | (1358.32) | (2952.51) |
| P-value, test of diff. | 0.000 | 0.410 | 0.285 | 0.161 | 0.157 | 0.120 |
| CCM, Minority | | 3611 | 6064 | 7421 | 9915 | 27038 |
| CCM, White | | 4243 | 6466 | 8299 | 11567 | 30803 |

Notes: N = 43,019 (37,653 minority youth including Black, Hispanic, Asian, and Mixed Race/Other, 5,366 White youth). 390 observations are dropped due to missing race/ethnicity. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.12: Employment and Earnings Effects by Race/Ethnicity, ITT

| | Y1 | Y2 | Y3 | Y4 | Cumulative | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|---|---|---|---|
| | Employment | | | | | Earnings, Winsorized at 99th Percentile | | | | |
| ITT, White | 0.0049 | -0.0007 | -0.0099 | -0.0073 | -0.0084 | -54.07 | -128.06 | -291.08 | -349.92 | -834.28 |
| | (0.0114) | (0.0122) | (0.0128) | (0.0129) | (0.0075) | (132.00) | (213.69) | (296.68) | (402.37) | (874.83) |
| ITT, Black | 0.0078 | 0.0044 | 0.0011 | -0.0099 | -0.0001 | 6.42 | 74.21 | 53.49 | 27.77 | 198.60 |
| | (0.0064) | (0.0063) | (0.0067) | (0.0067) | (0.0037) | (65.54) | (107.32) | (140.97) | (179.37) | (395.84) |
| ITT, Hispanic | 0.0165** | 0.0066 | 0.0198** | 0.0124 | 0.0093** | 203.13** | 213.19 | 283.35* | 380.19* | 1095.38** |
| | (0.0077) | (0.0077) | (0.0080) | (0.0081) | (0.0047) | (83.64) | (138.10) | (171.42) | (225.46) | (498.19) |
| ITT, Asian | 0.0252** | 0.0026 | -0.0079 | 0.0242* | 0.0063 | -2.12 | 61.98 | 500.69 | 1114.00** | 1811.90** |
| | (0.0121) | (0.0121) | (0.0126) | (0.0125) | (0.0073) | (108.48) | (200.40) | (309.80) | (441.67) | (884.08) |
| ITT, Other | 0.010 | 0.0338* | -0.0185 | -0.0074 | 0.0043 | 103.15 | 427.22 | -42.74 | 368.94 | 951.71 |
| | (0.0194) | (0.0192) | (0.0200) | (0.0204) | (0.0120) | (200.40) | (319.18) | (452.16) | (589.43) | (1269.50) |
| P-value, all equal | 0.671 | 0.64 | 0.117 | 0.073 | 0.296 | 0.318 | 0.565 | 0.321 | 0.099 | 0.157 |
| CM, White | 0.751 | 0.694 | 0.570 | 0.617 | 0.917 | 3662 | 5636 | 7202 | 10354 | 27023 |
| CM, Black | 0.715 | 0.744 | 0.676 | 0.708 | 0.931 | 3551 | 5917 | 7275 | 9191 | 25971 |
| CM, Hispanic | 0.686 | 0.718 | 0.656 | 0.684 | 0.916 | 3668 | 6325 | 7478 | 9766 | 27269 |
| CM, Asian | 0.643 | 0.675 | 0.614 | 0.650 | 0.914 | 2956 | 5195 | 7122 | 11301 | 26837 |
| CM, Other | 0.685 | 0.704 | 0.682 | 0.685 | 0.916 | 3512 | 5577 | 7630 | 9817 | 26577 |

Notes: N = 5,366 White, N = 17,636 Black, N = 12,427 Hispanic, N = 5,578 Asian, and N = 2,012 Mixed Race/Other. 390 observations dropped due to missing race/ethnicity. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means. P-value from test of null hypothesis that all treatment effects are equal across groups. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.13: Employment and Earnings Effects by Race/Ethnicity, IV

| | | Y1 | Y2 | Y3 | Y4 | Cumul. | Y1 | Y2 | Y3 | Y4 | Cumul. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | First Stage | Employment | | | | | Earnings, Winsorized at 99th Percentile | | | | |
| IV, White | 0.2973*** | 0.0162 | -0.0039 | -0.0333 | -0.0246 | -0.0283 | -191.62 | -453.07 | -978.84 | -1183.07 | -2861.66 |
| | (0.0088) | (0.0385) | (0.0412) | (0.0433) | (0.0437) | (0.0255) | (445.14) | (720.36) | (1002.05) | (1358.11) | (2953.12) |
| IV, Black | 0.4039*** | 0.0194 | 0.0108 | 0.0034 | -0.0244 | -0.0002 | 24.02 | 198.01 | 149.00 | 91.52 | 495.10 |
| | (0.0052) | (0.0157) | (0.0155) | (0.0165) | (0.0167) | (0.0092) | (162.06) | (265.35) | (348.52) | (443.42) | (979.76) |
| IV, Hispanic | 0.4152*** | 0.0396** | 0.0157 | 0.0475** | 0.0299 | 0.0224** | 486.56** | 508.28 | 676.96 | 907.86* | 2633.74** |
| | (0.0062) | (0.0186) | (0.0185) | (0.0193) | (0.0196) | (0.0114) | (201.36) | (332.35) | (412.47) | (542.62) | (1199.97) |
| IV, Asian | 0.4830*** | 0.0521** | 0.0055 | -0.0153 | 0.0500* | 0.0131 | 9.06 | 151.61 | 1064.39* | 2344.75** | 3758.13** |
| | (0.0094) | (0.0250) | (0.0250) | (0.0261) | (0.0259) | (0.0151) | (224.45) | (415.00) | (642.79) | (915.64) | (1833.46) |
| IV, Other | 0.3925*** | 0.0252 | 0.0857* | -0.0457 | -0.0189 | 0.011 | 271.80 | 1101.66 | -76.36 | 982.90 | 2391.00 |
| | (0.0155) | (0.0496) | (0.0493) | (0.0511) | (0.0523) | (0.0308) | (512.13) | (817.49) | (1156.50) | (1509.06) | (3251.01) |
| P-value, all equal | 0.000 | 0.802 | 0.649 | 0.133 | 0.078 | 0.340 | 0.350 | 0.606 | 0.395 | 0.134 | 0.222 |
| CCM, White | | 0.752 | 0.711 | 0.622 | 0.645 | 0.951 | 4242 | 6465 | 8300 | 11564 | 30799 |
| CCM, Black | | 0.723 | 0.764 | 0.699 | 0.74 | 0.938 | 3736 | 6296 | 7765 | 9704 | 27520 |
| CCM, Hisp. | | 0.680 | 0.722 | 0.637 | 0.685 | 0.91 | 3681 | 6357 | 7439 | 10001 | 27458 |
| CCM, Asian | | 0.629 | 0.667 | 0.617 | 0.626 | 0.899 | 3077 | 5128 | 6197 | 10307 | 24832 |
| CCM, Other | | 0.692 | 0.692 | 0.710 | 0.704 | 0.918 | 3876 | 5324 | 8405 | 9888 | 27502 |

Notes: N = 5,366 White, N = 17,636 Black, N = 12,427 Hispanic, N = 5,578 Asian, and N = 2,012 Mixed Race/Other. 390 observations dropped due to missing race/ethnicity. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CCM shows control complier means. P-value from test of null hypothesis that all treatment effects are equal across groups. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.14: Employment and Earnings Effects by Gender

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | | | Employment | | |
| ITT, Male | | 0.0044 | 0.0098 | 0.0054 | 0.0016 | 0.0028 |
| | | (0.0066) | (0.0067) | (0.0069) | (0.0070) | (0.0044) |
| ITT, Female | | 0.0190*** | 0.0029 | 0.0011 | 0.0008 | 0.0031 |
| | | (0.0052) | (0.0052) | (0.0056) | (0.0056) | (0.0029) |
| P-value, test of diff. | | 0.083 | 0.409 | 0.629 | 0.935 | 0.964 |
| CM, Male | | 0.658 | 0.659 | 0.585 | 0.615 | 0.894 |
| CM, Female | | 0.733 | 0.766 | 0.699 | 0.731 | 0.943 |
| IV, Male | 0.3962*** | 0.0111 | 0.0249 | 0.014 | 0.0039 | 0.007 |
| | (0.0051) | (0.0166) | (0.0168) | (0.0173) | (0.0177) | (0.0111) |
| IV, Female | 0.4106*** | 0.0462*** | 0.007 | 0.0032 | 0.002 | 0.0071 |
| | (0.0044) | (0.0128) | (0.0127) | (0.0136) | (0.0135) | (0.0070) |
| P-value, test of diff. | 0.031 | 0.094 | 0.396 | 0.62 | 0.931 | 0.993 |
| CCM, Male | | 0.675 | 0.666 | 0.601 | 0.633 | 0.898 |
| CCM, Female | | 0.713 | 0.773 | 0.706 | 0.742 | 0.942 |
| | | | | Earnings, Winsorized at 99th Percentile | | |
| ITT, Male | | 50.05 | 155.43 | 155.04 | 246.22 | 637.00 |
| | | (62.35) | (106.86) | (144.91) | (194.84) | (418.06) |
| ITT, Female | | 63.87 | 66.31 | 108.40 | 191.24 | 475.58 |
| | | (59.27) | (96.98) | (129.47) | (170.58) | (370.00) |
| P-value, test of diff. | | 0.872 | 0.537 | 0.810 | 0.832 | 0.772 |
| CM, Male | | 2968 | 4963 | 6416 | 8675 | 23111 |
| CM, Female | | 3952 | 6642 | 8096 | 10861 | 29640 |
| | First Stage | | | | | |
| IV, Male | 0.3962*** | 133.37 | 401.58 | 411.13 | 636.41 | 1617.41 |
| | (0.0051) | (157.18) | (269.46) | (365.11) | (491.50) | (1055.15) |
| IV, Female | 0.4106*** | 160.27 | 170.37 | 272.34 | 481.08 | 1155.60 |
| | (0.0044) | (144.31) | (236.12) | (315.34) | (415.27) | (901.22) |
| P-value, test of diff. | 0.031 | 0.900 | 0.519 | 0.773 | 0.809 | 0.739 |
| CCM, Male | | 3195 | 5161 | 6676 | 8963 | 24054 |
| CCM, Female | | 4033 | 6830 | 8186 | 11157 | 30255 |

Notes: Notes: N = 43,409 (18,539 male youth, 24,870 female youth). Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.15: Employment and Earnings Effects by Neighborhood: Above/Below Median in Opportunity Insights Upward Mobility Ranking

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | | | Employment | | |
| ITT, Below Median | | 0.0144** | 0.0098* | 0.0079 | 0.0016 | 0.0036 |
| | | (0.0058) | (0.0057) | (0.0060) | (0.0061) | (0.0035) |
| ITT, Above Median | | 0.0112* | 0.0018 | -0.0021 | 0.0006 | 0.0023 |
| | | (0.0058) | (0.0059) | (0.0062) | (0.0063) | (0.0036) |
| P-value, test of diff. | | 0.699 | 0.332 | 0.25 | 0.907 | 0.799 |
| CM, Below Median | | 0.696 | 0.729 | 0.660 | 0.695 | 0.924 |
| CM, Above Median | | 0.706 | 0.711 | 0.640 | 0.669 | 0.921 |
| IV, Below Median | 0.4182*** | 0.0345** | 0.0235* | 0.0192 | 0.0036 | 0.0079 |
| | (0.0047) | (0.0138) | (0.0137) | (0.0144) | (0.0146) | (0.0083) |
| IV, Above Median | 0.3903*** | 0.0288* | 0.0047 | -0.0046 | 0.0019 | 0.0063 |
| | (0.0047) | (0.0150) | (0.0152) | (0.0159) | (0.0161) | (0.0092) |
| P-value, test of diff. | 0.000 | 0.781 | 0.358 | 0.266 | 0.937 | 0.896 |
| CCM, Below Median | | 0.700 | 0.741 | 0.672 | 0.711 | 0.928 |
| CCM, Above Median | | 0.693 | 0.714 | 0.651 | 0.681 | 0.919 |
| | | | | Earnings, Winsorized at 99th Percentile | | |
| ITT, Below Median | | 54.02 | 151.67 | 64.28 | 83.12 | 371.79 |
| | | (60.66) | (99.13) | (127.17) | (164.10) | (362.56) |
| ITT, Above Median | | 62.04 | 57.01 | 194.70 | 352.70* | 722.90* |
| | | (61.45) | (104.33) | (145.73) | (197.76) | (420.14) |
| P-value, test of diff. | | 0.926 | 0.511 | 0.500 | 0.294 | 0.527 |
| CM, Below Median | | 3587 | 6006 | 7239 | 9270 | 26141 |
| CM, Above Median | | 3476 | 5844 | 7518 | 10591 | 27570 |
| | First Stage | | | | | |
| IV, Below Median | 0.4182*** | 133.43 | 369.80 | 161.28 | 210.11 | 889.18 |
| | (0.0047) | (144.87) | (236.69) | (303.70) | (391.74) | (866.46) |
| IV, Above Median | 0.3903*** | 167.16 | 157.30 | 520.99 | 927.76* | 1873.76* |
| | (0.0047) | (157.48) | (267.32) | (373.30) | (506.98) | (1077.22) |
| P-value, test of diff. | 0.000 | 0.875 | 0.552 | 0.455 | 0.263 | 0.476 |
| CCM, Below Median | | 3774 | 6289 | 7762 | 10028 | 27882 |
| CCM, Above Median | | 3581 | 5957 | 7321 | 10457 | 27397 |

Notes: N = 43,408 (21,860 below median, 21,548 above median, 1 observation missing zip code). Uses within-sample median of Opportunity Insights "upward mobility" index: the average percentile rank for children whose parents were in the 25th percentile of the national income distribution. We map Census tract-level data onto participant zip code, see text for details. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.16: Letter Information and Application Behavior for Treatment Group by Subgroup

| | Has Letter | Average Rating | Applied to Our Job | Submitted Letter |
|---|---|---|---|---|
| White | 0.296 | 6.09 | 0.073 | 0.222 |
| Non-White | 0.420 | 5.66 | 0.083 | 0.158 |
| Black | 0.404 | 5.54 | 0.087 | 0.167 |
| Hispanic | 0.416 | 5.68 | 0.071 | 0.130 |
| Asian | 0.483 | 5.85 | 0.121 | 0.200 |
| Male | 0.396 | 5.62 | 0.077 | 0.162 |
| Female | 0.410 | 5.75 | 0.086 | 0.167 |
| In HS Sample | 0.412 | 5.61 | 0.082 | 0.230 |
| Not in HS Sample | 0.398 | 5.77 | 0.082 | 0.111 |
| Under 18 | 0.403 | 5.64 | 0.084 | 0.226 |
| 18 and Over | 0.407 | 5.79 | 0.079 | 0.052 |
| Above Median in OI Rank | 0.390 | 5.80 | 0.077 | 0.167 |
| Below Median in OI Rank | 0.418 | 5.60 | 0.086 | 0.163 |
| Never Employed Pre-SYEP | 0.395 | 5.61 | 0.077 | 0.214 |
| Ever Employed Pre-SYEP | 0.412 | 5.77 | 0.086 | 0.128 |
| High Rating | 0.753 | 6.04 | 0.093 | 0.250 |
| Low Rating | 0.307 | 3.92 | 0.074 | 0.167 |

Notes: Means shown for treatment group only, N = 21,714 (except for high/low rating, which is limited to those with a rating, N = 14,400). Average rating conditional on being sent a letter, N = 8,780; application probability conditional on being invited to apply, N = 2,000 (1,346 for rating categories); and submission probability conditional on applying, N = 164 (116 for rating categories). Median OI Rank is the within-sample median of the Opportunity Insights "upward mobility" percentile rank. All differences in having a letter and in average ratings between two groups (i.e., White/Minority, Male/Female, High School/Not in HS, Under/Over 18, Above/Below median OI rank, Never/Ever Employed Pre-SYEP, and High/Low Ratings) are statistically different except for having a letter between those under and over 18. None of the differences in application or letter submission rates are significantly different except for the high school and age differences in submitting the letter.

Table A.17: Education Descriptive Statistics

| | N | Control Mean | Treatment Mean | Test of Difference |
|---|---|---|---|---|
| | Education Sample | | | |
| Age | 19714 | 15.96 | 15.95 | 0.357 |
| Male | 19714 | 0.452 | 0.445 | 0.344 |
| Black | 19656 | 0.426 | 0.424 | 0.854 |
| Hispanic | 19656 | 0.309 | 0.307 | 0.821 |
| Asian | 19656 | 0.139 | 0.137 | 0.794 |
| White | 19656 | 0.074 | 0.084 | 0.009 |
| Grade Level | 19714 | 10.04 | 10.03 | 0.344 |
| Share Enrolled Days Present | 19714 | 0.902 | 0.899 | 0.169 |
| Missing GPA | 19714 | 0.100 | 0.101 | 0.848 |
| GPA (100 point scale) | 17732 | 79.73 | 79.34 | 0.033 |
| In College Sample | 19714 | 0.903 | 0.903 | 0.998 |
| Not in UI Data | 19714 | 0.008 | 0.010 | 0.411 |
| Never Employed Pre-SYEP | 19714 | 0.614 | 0.621 | 0.284 |
| Ever Worked, Year -4 | 19714 | 0.041 | 0.040 | 0.688 |
| Earnings, Year -4 | 19714 | 64.93 | 81.84 | 0.247 |
| Ever Worked, Year -3 | 19714 | 0.134 | 0.134 | 0.916 |
| Earnings, Year -3 | 19714 | 169.01 | 180.68 | 0.497 |
| Ever Worked, Year -2 | 19714 | 0.305 | 0.304 | 0.889 |
| Earnings, Year -2 | 19714 | 411.70 | 402.85 | 0.629 |
| Ever Worked, Year -1 | 19714 | 0.958 | 0.960 | 0.565 |
| Earnings, Year -1 | 19714 | 1545.18 | 1534.16 | 0.604 |
| Joint F-test | F(37, 19063) = 1.242, p=.149 | | | |

Notes: Table shows non-missing summary statistics for the expected in high school sample (see text for details). Test of difference reports the p-value from a regression of each characteristic on a treatment indicator, controlling for a cohort indicator and using standard errors clustered on individual.

Table A.18: Joint Employment and School Attainment Outcomes

| | Panel A: On-Time Graduation | | | |
|---|---|---|---|---|
| | Ever Work, On-time Grad | Never Work, On-time Grad | Ever Work, Not On-time | Never Work, Not On-time |
| ITT | -0.0044 | -0.0018 | 0.0096** | -0.0032 |
| | (0.0050) | (0.0030) | (0.0042) | (0.0026) |
| CM | 0.736 | 0.049 | 0.177 | 0.037 |
| Sent Letter (IV) | -0.0101 | -0.0052 | 0.0231** | -0.0075 |
| | (0.0120) | (0.0073) | (0.0101) | (0.0062) |
| CCM | 0.777 | 0.050 | 0.139 | 0.034 |
| | Panel B: Any Graduation | | | |
| | Ever Work, Graduated | Never Work, Graduated | Ever Work, Not Graduated | Never Work, Not Graduated |
| ITT | 0.0022 | -0.0028 | 0.0029 | -0.002 |
| | (0.0048) | (0.0031) | (0.0039) | (0.0025) |
| CM | 0.781 | 0.053 | 0.132 | 0.034 |
| Sent Letter (IV) | 0.0052 | -0.0077 | 0.0072 | -0.0049 |
| | (0.0117) | (0.0075) | (0.0094) | (0.0060) |
| CCM | 0.811 | 0.054 | 0.106 | 0.03 |
| | Panel C: Any Graduation or Continued Attendance | | | |
| | Ever Work, Persisted | Never Work, Persisted | Ever Work, Not Persisted | Never Work, Not Persisted |
| ITT | -0.0012 | -0.0021 | 0.0064* | -0.0025 |
| | (0.0049) | (0.0032) | (0.0038) | (0.0024) |
| CM | 0.795 | 0.055 | 0.118 | 0.031 |
| Sent Letter (IV) | -0.0028 | -0.0063 | 0.0154* | -0.0062 |
| | (0.0117) | (0.0077) | (0.0092) | (0.0057) |
| CCM | 0.827 | 0.056 | 0.089 | 0.028 |

Notes: N = 19,714. Analysis conducted on the main education sample (non-charter 8th–12th graders in the pre-randomization year, see text for details). First stage for this subsample is 0.413. Panel A shows whether someone ever worked during the 4-year follow up and whether they graduated on-time (i.e., 4th-year graduation). Panel B shows whether someone ever worked during the 4-year follow up and whether they ever graduated (i.e., 4th-, 5th-, or 6th-year graduation). Panel C shows whether someone ever worked during the 4-year follow up and whether they either graduated or had positive days attended in the last year of our data. CCM shows control complier means, which may not total to 1 across categories due to estimation error in the IV and the inclusion of different sets of covariates in the post double-selection LASSO. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.19: Education Effects by GPA

| | Ever Enrolled Y1 | % Enrolled Days Present Y1 | GPA Y1 | Credits Attempted Y1-4 | % Credits Earned Y1-4 | Graduated On-Time | Ever Graduated | Graduated or Still Attending | On-time College |
|---|---|---|---|---|---|---|---|---|---|
| IV, Below | -0.0247* | -0.0083 | -0.8497* | 0.1626 | -0.002 | -0.0569** | -0.0175 | -0.0232 | -0.0329 |
| Median GPA | (0.0148) | (0.0141) | (0.4957) | (0.4681) | (0.0164) | (0.0232) | (0.0216) | (0.0212) | (0.0264) |
| IV, Above | -0.0016 | 0.001 | -0.3397 | -0.0176 | -0.0038 | -0.001 | 0.0014 | 0.0019 | 0.0004 |
| Median GPA | (0.0049) | (0.0054) | (0.3136) | (0.1984) | (0.0063) | (0.0082) | (0.0074) | (0.0072) | (0.0148) |
| P, test of diff. | 0.141 | 0.54 | 0.385 | 0.724 | 0.921 | 0.023 | 0.407 | 0.265 | 0.271 |
| CCM, Below | 0.945 | 0.772 | 72.984 | 18.914 | 0.776 | 0.753 | 0.83 | 0.845 | 0.543 |
| CCM, Above | 0.993 | 0.937 | 89.624 | 18.109 | 0.976 | 0.974 | 0.977 | 0.978 | 0.885 |

Notes: N = 17,732, all in main education sample who are not missing baseline GPA (8,868 above median, 8,864 below). Median GPA cut-off is 80.85. Credits attempted and % credits earned equal 0 for those not in school. On-time graduation equals 1 for public school, non-charter students who graduate within 4 years and do not transfer to GED programs or other districts. Ever graduated adds any 5th- and 6th-year graduation observed during the follow-up period. Graduated or still attending equals 1 if student either graduated or has positive days attended in most recent academic year. College enrollment is only measured within 6 months after a student's on-time graduation date, regardless of graduation status. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are shown in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.20: Employment and Earnings Effects by GPA

| | First Stage | Y1 | Y2 | Y3 | Y4 | Cumulative |
|---|---|---|---|---|---|---|
| | | Panel A: Employment | | | | |
| IV, Below Median | 0.3704*** | 0.0642** | 0.0409 | 0.0409 | 0.0099 | 0.025 |
| GPA | (0.0072) | (0.0260) | (0.0253) | (0.0260) | (0.0264) | (0.0153) |
| IV, Above Median | 0.4615*** | -0.0107 | 0.0184 | -0.0141 | 0.0261 | -0.0033 |
| GPA | (0.0076) | (0.0208) | (0.0205) | (0.0215) | (0.0208) | (0.0118) |
| P, test of diff. | 0.000 | 0.025 | 0.49 | 0.104 | 0.632 | 0.143 |
| CCM, Below | | 0.64 | 0.705 | 0.654 | 0.696 | 0.919 |
| CCM, Above | | 0.693 | 0.696 | 0.638 | 0.684 | 0.939 |
| | | Panel B: Earnings | | | | |
| IV, Below Median | | -43.66 | 188.02 | 544.23 | 537.75 | 807.55 |
| GPA | | (195.96) | (365.87) | (444.48) | (605.06) | (1348.27) |
| IV, Above Median | | 28.90 | 16.99 | 238.30 | 1055.54** | 1234.33 |
| GPA | | (141.89) | (244.72) | (336.28) | (478.33) | (958.11) |
| P, test of diff. | | 0.765 | 0.699 | 0.584 | 0.503 | 0.797 |
| CCM, Below | | 2432 | 4583 | 5961 | 7688 | 21084 |
| CCM, Above | | 2405 | 4210 | 4795 | 6653 | 18167 |
| | | Panel C: Earnings, Winsorized at 99th Percentile | | | | |
| IV, Below Median | | -34.81 | 317.96 | 529.69 | 442.99 | 1304.21 |
| GPA | | (171.32) | (327.38) | (436.42) | (589.36) | (1202.26) |
| IV, Above Median | | 51.63 | 28.74 | 242.02 | 965.64** | 1240.49 |
| GPA | | (129.78) | (237.55) | (328.22) | (460.33) | (929.01) |
| P, test of diff. | | 0.687 | 0.475 | 0.599 | 0.486 | 0.967 |
| CCM, Below | | 2422 | 4436 | 5970 | 7732 | 20529 |
| CCM, Above | | 2367 | 4198 | 4775 | 6700 | 18118 |
| | | Panel D: Number of Quarters Worked | | | | |
| IV, Below Median | | 0.042 | 0.132* | 0.104 | 0.025 | 0.307 |
| GPA | | (0.073) | (0.079) | (0.085) | (0.091) | (0.230) |
| IV, Above Median | | 0.028 | -0.004 | 0.021 | 0.054 | 0.097 |
| GPA | | (0.058) | (0.065) | (0.071) | (0.074) | (0.189) |
| P, test of diff. | | 0.883 | 0.183 | 0.454 | 0.805 | 0.481 |
| CCM, Below | | 1.44 | 1.65 | 1.88 | 1.95 | 6.92 |
| CCM, Above | | 1.43 | 1.71 | 1.71 | 1.93 | 6.78 |

Notes: N = 17,732, all in main education sample who are not missing baseline GPA (8,868 above median, 8,864 below). Median GPA cut-off is 80.85. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.21: Joint Employment and School Attainment Outcomes by GPA, IV

| | Panel A: On-Time Graduation | | | |
|---|---|---|---|---|
| | Ever Work, On-time Grad | Never Work, On-time Grad | Ever Work, Not On-time | Never Work, Not On-time |
| IV, Below Median GPA | -0.0404* | -0.0169* | 0.0643*** | -0.0094 |
| | (0.0241) | (0.0101) | (0.0231) | (0.0120) |
| IV, Above Median GPA | -0.0051 | 0.0048 | 0.0034 | -0.0019 |
| | (0.0136) | (0.0112) | (0.0072) | (0.0042) |
| P, test of diff. | 0.202 | 0.149 | 0.012 | 0.557 |
| CCM, Below | 0.706 | 0.047 | 0.214 | 0.035 |
| CCM, Above | 0.922 | 0.052 | 0.016 | 0.009 |
| | Panel B: Any Graduation | | | |
| | Ever Work, Graduated | Never Work, Graduated | Ever Work, Not Graduated | Never Work, Not Graduated |
| IV, Below Median GPA | 0.0058 | -0.0229** | 0.0191 | -0.0023 |
| | (0.0229) | (0.0109) | (0.0206) | (0.0113) |
| IV, Above Median GPA | -0.0041 | 0.0049 | 0.0007 | -0.0019 |
| | (0.0132) | (0.0112) | (0.0063) | (0.0042) |
| P, test of diff. | 0.707 | 0.076 | 0.394 | 0.971 |
| CCM, Below | 0.773 | 0.057 | 0.146 | 0.025 |
| CCM, Above | 0.926 | 0.052 | 0.014 | 0.009 |
| | Panel C: Any Graduation or Continued Attendance | | | |
| | Ever Work, Persisted | Never Work, Persisted | Ever Work, Not Persisted | Never Work, Not Persisted |
| IV, Below Median GPA | -0.0009 | -0.0220* | 0.0256 | -0.0033 |
| | (0.0227) | (0.0112) | (0.0203) | (0.0110) |
| IV, Above Median GPA | -0.0051 | 0.0065 | 0.0012 | -0.0036 |
| | (0.0131) | (0.0112) | (0.0060) | (0.0041) |
| P, test of diff. | 0.872 | 0.074 | 0.251 | 0.979 |
| CCM, Below | 0.786 | 0.058 | 0.133 | 0.023 |
| CCM, Above | 0.928 | 0.051 | 0.012 | 0.01 |

Notes: N=17,732, all in main education sample who are not missing baseline GPA (8,868 above median, 8,864 below). Median GPA = 80.85. First stage for below median GPA = 0.370, for above median GPA = 0.462. Panel A shows whether someone ever worked during the 4-year follow up and whether they graduated on-time (i.e., 4th-year graduation). Panel B shows whether someone ever worked during the 4-year follow up and whether they ever graduated (i.e., 4th-, 5th-, or 6th-year graduation). Panel C shows whether someone ever worked during the 4-year follow up and whether they either graduated or had positive days attended in the last year of our data. CCM shows control complier means, which may not total to 1 across categories due to estimation error in the IV and the inclusion of different sets of covariates in the post double-selection LASSO. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A.22: Labor Market Effects, Alternative Covariates

| Year | 1 | 2 | 3 | 4 | Cumulative |
|---|---|---|---|---|---|
| | Panel A: No Covariates | | | | |
| | Employment | | | | |
| ITT | 0.0117*** | 0.0049 | 0.0026 | 0.0009 | 0.0025 |
| | (0.0044) | (0.0043) | (0.0045) | (0.0045) | (0.0026) |
| CM | 0.701 | 0.72 | 0.65 | 0.682 | 0.922 |
| Sent Letter (IV) | 0.0289*** | 0.0122 | 0.0065 | 0.0023 | 0.0062 |
| | (0.0108) | (0.0106) | (0.0113) | (0.0110) | (0.0063) |
| CCM | 0.7 | 0.73 | 0.663 | 0.697 | 0.925 |
| | Earnings, Winsorized at 99th Percentile | | | | |
| ITT | 44.12 | 100.2 | 138.52 | 243.75* | 547.88* |
| | (51.84) | (81.53) | (106.39) | (138.57) | (320.38) |
| CM | 3532 | 5925 | 7378 | 9927 | 26852 |
| Sent Letter (IV) | 109.11 | 247.81 | 342.56 | 602.82* | 1354.95* |
| | (128.19) | (201.58) | (263.07) | (342.60) | (792.06) |
| CCM | 3722 | 6151 | 7542 | 10183 | 27655 |
| | Panel B: All Covariates | | | | |
| | Employment | | | | |
| ITT | 0.0125*** | 0.0058 | 0.003 | 0.0012 | 0.0028 |
| | (0.0041) | (0.0041) | (0.0043) | (0.0044) | (0.0025) |
| CM | 0.701 | 0.720 | 0.65 | 0.682 | 0.922 |
| Sent Letter (IV) | 0.0309*** | 0.0144 | 0.0074 | 0.0029 | 0.0069 |
| | (0.0101) | (0.0102) | (0.0107) | (0.0108) | (0.0062) |
| CCM | 0.698 | 0.728 | 0.662 | 0.696 | 0.924 |
| | Earnings, Winsorized at 99th Percentile | | | | |
| ITT | 53.04 | 103.78 | 131.62 | 219.40* | 528.91* |
| | (43.02) | (71.82) | (96.34) | (128.09) | (276.64) |
| CM | 3532 | 5925 | 7378 | 9927 | 26852 |
| Sent Letter (IV) | 131.15 | 256.62 | 325.45 | 542.50* | 1307.82* |
| | (106.32) | (177.53) | (238.17) | (316.64) | (683.88) |
| CCM | 3700 | 6143 | 7559 | 10243 | 27702 |

Notes: N = 43,409. Panel A shows results with no coviarates other than cohort indicator. Panel B uses all available covariates (see text) rather than post-double-selection LASSO-selected covariates that are used in the main results. Winsorization in Panel B recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.23: Education Effects, Alternative Covariates

| | Ever Enrolled Y1 | % Enrolled Days Present | GPA Y1 | Credits Attempted Y1-4 | % Credits Earned Y1-4 | Graduated On-Time | Ever Graduated | Graduated or Still Attending | On-time College |
|---|---|---|---|---|---|---|---|---|---|
| | Panel A: No Covariates | | | | | | | | |
| ITT | -0.003 | -0.004 | -0.463** | 0.187 | -0.002 | -0.014** | -0.006 | -0.009* | -0.015** |
| | (0.003) | (0.004) | (0.184) | (0.144) | (0.004) | (0.006) | (0.005) | (0.005) | (0.007) |
| CM | 0.946 | 0.829 | 80.128 | 18.958 | 0.818 | 0.785 | 0.834 | 0.851 | 0.672 |
| Sent Letter (IV) | -0.008 | -0.009 | -1.113** | 0.452 | -0.006 | -0.034** | -0.015 | -0.022* | -0.035** |
| | (0.008) | (0.009) | (0.445) | (0.350) | (0.011) | (0.014) | (0.013) | (0.012) | (0.017) |
| CCM | 0.961 | 0.861 | 82.556 | 18.233 | 0.856 | 0.846 | 0.877 | 0.895 | 0.742 |
| | Panel B: All Covariates | | | | | | | | |
| ITT | -0.002 | 0.001 | -0.135 | 0.076 | 0.002 | -0.007* | -0.001 | -0.004 | -0.005 |
| | (0.003) | (0.003) | (0.098) | (0.100) | (0.003) | (0.004) | (0.004) | (0.004) | (0.006) |
| CM | 0.946 | 0.829 | 80.128 | 18.958 | 0.818 | 0.785 | 0.834 | 0.851 | 0.672 |
| Sent Letter (IV) | -0.004 | 0.004 | -0.324 | 0.183 | 0.006 | -0.017* | -0.002 | -0.01 | -0.011 |
| | (0.007) | (0.007) | (0.235) | (0.241) | (0.007) | (0.010) | (0.010) | (0.010) | (0.013) |
| CCM | 0.957 | 0.848 | 81.767 | 18.502 | 0.844 | 0.828 | 0.864 | 0.883 | 0.718 |
| N | 19714 | 19714 | 18237 | 19714 | 19714 | 19714 | 19714 | 19714 | 17810 |

Notes: Analysis conducted on main education sample. Panel A shows results with no coviarates other than cohort indicator. Panel B uses all available covariates (see text) rather than post-double-selection LASSO-selected covariates that are used in the main results. Credits attempted and earned equal 0 for those not in school. On-time graduation equals 1 for public school, non-charter students who graduate within 4 years and do not transfer to GED programs or other districts. Ever graduated adds any 5th- and 6th-year graduation observed during the follow-up period. Graduated or still attending equals 1 if student either graduated or has positive days attended in most recent academic year. College enrollment is only measured within 6 months after a student's on-time graduation date, regardless of graduation status. CM shows control means; CCM shows control complier means. Standard errors clustered on individual are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Table A.24: Baseline Characteristics, Unopened Surveys versus Main Control Group

| | Unopened Surveys | Control | Test of Difference |
|---|---|---|---|
| N | 25813 | 21695 | |
| Age | 17.24 | 17.17 | 0.002 |
| Male | 0.427 | 0.427 | 0.894 |
| Black | 0.437 | 0.409 | 0.000 |
| Hispanic | 0.246 | 0.289 | 0.000 |
| Asian | 0.082 | 0.129 | 0.000 |
| White | 0.188 | 0.124 | 0.000 |
| Other Race | 0.047 | 0.049 | 0.746 |
| In High School | 0.746 | 0.755 | 0.014 |
| HS Graduate | 0.050 | 0.044 | 0.003 |
| In College | 0.174 | 0.173 | 0.677 |
| Not in UI Data | 0.008 | 0.006 | 0.078 |
| Never Employed Pre-SYEP | 0.429 | 0.450 | 0.000 |
| Ever Worked, Year -4 | 0.170 | 0.153 | 0.000 |
| Earnings, Year -4 | 322 | 303 | 0.203 |
| Ever Worked, Year -3 | 0.293 | 0.266 | 0.000 |
| Earnings, Year -3 | 609 | 574 | 0.091 |
| Ever Worked, Year -2 | 0.459 | 0.437 | 0.000 |
| Earnings, Year -2 | 1093 | 1052 | 0.146 |
| Ever Worked, Year -1 | 0.962 | 0.965 | 0.062 |
| Earnings, Year -1 | 2331 | 2334 | 0.719 |
| No Education Match | 0.185 | 0.126 | 0.000 |
| In HS Sample | 0.409 | 0.454 | 0.000 |
| Joint F-test | F(24, 45597) = 35.492, p=0 | | |

Notes: Table tests difference of means between all youth in unopened surveys (excluded from our main sample) and our control group (on an opened survey). Test of difference controls for cohort indicator and uses cluster-robust standard errors. 496 youth are missing race/ethnicity.

Table A.25: Outcomes, Unopened Surveys versus Main Control Group

| | Unopened Surveys | Control | Test of Difference |
|---|---|---|---|
| | Panel A: Labor Market Outcomes | | |
| N | 25813 | 21695 | |
| Employment Y1 | 0.715 | 0.701 | 0.000 |
| Employment Y2 | 0.715 | 0.720 | 0.278 |
| Employment Y3 | 0.640 | 0.650 | 0.143 |
| Employment Y4 | 0.666 | 0.682 | 0.000 |
| Employment Cumulative | 0.919 | 0.922 | 0.210 |
| Earnings Y1 | 3617 | 3532 | 0.135 |
| Earnings Y2 | 6157 | 5925 | 0.006 |
| Earnings Y3 | 7561 | 7378 | 0.032 |
| Earnings Y4 | 10050 | 9927 | 0.332 |
| Earnings Cumulative | 27500 | 26852 | 0.031 |
| Joint F-test, Employment Outcomes | F(8, 45597) = 6.482, p=0 | | |
| | Panel B: Education Outcomes | | |
| N | 10564 | 9857 | |
| Enrolled Y1 | 0.934 | 0.946 | 0.000 |
| Perc. Days Present Y1 | 0.808 | 0.829 | 0.000 |
| GPA Y1 | 79.03 | 80.13 | 0.000 |
| Credit Attempted Y1-4 | 18.68 | 18.96 | 0.054 |
| Perc. Credits Earned Y1-4 | 0.795 | 0.817 | 0.000 |
| Graduated On-time | 0.751 | 0.785 | 0.000 |
| Ever Graduated | 0.801 | 0.834 | 0.000 |
| Graduated or Still Attending | 0.817 | 0.851 | 0.000 |
| On-time College | 0.627 | 0.666 | 0.000 |
| On-time College | 0.635 | 0.672 | 0.000 |
| Joint F-test, Education Outcomes | F(6, 18093) = 10.185, p=0 | | |
| | Panel C: Job Application Outcomes | | |
| N | 636 | 2000 | |
| Clicked Link | 0.090 | 0.104 | 0.294 |
| Started Application | 0.075 | 0.089 | 0.288 |
| Uploaded Any File | 0.047 | 0.053 | 0.586 |
| Included Letter of Rec | 0.003 | 0.004 | 0.745 |
| Checked Selective Box | 0.039 | 0.053 | 0.137 |
| Joint F-test, Job App Outcomes | F(5, 2630) = .822, p=.534 | | |

Notes: Table tests difference of means between all youth in unopened surveys (excluded from our main sample) and our control group (on an opened survey), separately for employment outcomes and subset of youth in education sample. N = 18,396 for college test. To avoid using the smallest available sample and highly correlated outcomes for joint F-test, the labor market test excludes cumulative outcomes and the education joint test includes 5 high school outcomes and the indicator for graduating or still attending. Test of difference controls for cohort indicator and uses cluster-robust standard errors.

Table A.26: Employment and Earnings Effects, On Any Survey

| Year | 1 | 2 | 3 | 4 | Cumulative |
|---|---|---|---|---|---|
| | Panel A: Employment | | | | |
| ITT | 0.0063* | 0.0033 | 0.0018 | 0.0018 | -0.0003 |
| | (0.0032) | (0.0033) | (0.0034) | (0.0035) | (0.0020) |
| CM | 0.707 | 0.719 | 0.647 | 0.675 | 0.922 |
| Sent Letter (IV) | 0.0246* | 0.0131 | 0.0081 | 0.0071 | -0.0015 |
| | (0.0128) | (0.0129) | (0.0135) | (0.0137) | (0.0079) |
| CCM | 0.704 | 0.73 | 0.662 | 0.692 | 0.932 |
| | Panel B: Earnings | | | | |
| ITT | 25.24 | 65.20 | 122.14 | 152.78 | 362.76 |
| | (39.80) | (62.78) | (82.37) | (110.42) | (240.23) |
| CM | 3635 | 6102 | 7529 | 10124 | 27408 |
| Sent Letter (IV) | 111.34 | 256.36 | 505.73 | 626.43 | 1426.84 |
| | (156.53) | (247.02) | (324.06) | (434.43) | (945.49) |
| CCM | 3796 | 6232 | 7472 | 10301 | 27874 |
| | Panel C: Earnings, Winsorized at 99th Percentile | | | | |
| ITT | 15.67 | 58.78 | 113.44 | 174.43* | 373.51* |
| | (34.50) | (58.03) | (77.50) | (102.44) | (222.77) |
| CM | 3574 | 6017 | 7430 | 9954 | 27075 |
| Sent Letter (IV) | 63.01 | 231.26 | 447.05 | 681.39* | 1470.27* |
| | (135.76) | (228.33) | (305.00) | (403.16) | (876.71) |
| CCM | 3768 | 6168 | 7438 | 10104 | 27539 |

Notes: N = 69,222. Sample includes all youth on any survey, regardless of whether any supervisor opened the survey. Earnings winsorization recodes each quarter's highest earnings to the 99th percentile of all quarterly earnings before summing across years. CM shows control means; CCM shows control complier means. Regressions include baseline covariates. Standard errors clustered on individual are in parentheses. * p<0.1, ** p<0.05, *** p<0.01