Dynarski, S., Jacob, B., Kreisman, D. ''How Important Are Fixed Effects and Time Trends in Estimating Returns to Schooling? Evidence From A Replication of Jacobson, Lalonde and Sullivan, 2005''. (MS 10947).

The data for this project are subject to restricted access resulting from a research relationship between the University of Michigan and several unnamed community colleges. For this reason, the data cannot be made publicly available. For inquires about data access, please contact the University of Michigan Education Policy Initiative. For specific questions concerning the code or content, please contact Daniel Kreisman (dkreisman@gsu.edu).

We estimate 6 regressions in the main table (Table 3). We re-estimate the last 3 of these (columns 4-6 of Table 3) again by gender in Table 4. Then in the final table (Table 5), we re-estimate the last column of Table 3 for reference, and add two additional specifications.

Below we describe the sample limitation and give definitions of all the variables used in the regression models in the paper. These same variables are used the summary statistics table so we do not describe them again.

0.1 Sample Limitation

Variable id_match is the unique person identifier in the data.

- Drop observations after quarter 2, 2011. Variable quarter is defined by year-quarter (for example, 20112 is q2, 2011). Variable quarter_tq is same in Stata %tq time.
- Keep if indicator ccfirst$= 1$ - if first observed college enrollment is in our of our 5 community colleges.
- Keep those enrolled for the first time between 2002 and 2007. Variable first_term_enrolled is defined where terms are year-term; terms are defined over (3=spring, 5=summer, 7=fall, 9=winter; e.g. 20057 is fall, 2005).
- Drop observations where age or gender is missing.
- Keep those first enrolled between ages 21 and 45 (age_first_att_credit is age first attempting a credit).
- Limit sample to observations when individuals are between ages 17 and 65.
- Limit to those not going on to a 4-year school (ever_4year $= 1$ if ever enrolled in a 4-year school, and (hdc_ever is highest degree ever which = 4 if this is a BA).

0.2 Regression Tables

We crease several sets of globals to enter the tables that correspond to the indicators at the bottom of Table 3. We define each in turn here.

1. Earnings measures and enrollment:

   - Dependent variable is ui_quarterly_earnings_real, which are real quarterly earnings from Unemployment Insurance records from the Michigan Department of Licensing and Regulatory Affairs.
   - On their recommendation, values below $10 were recoded to $0.

2. Main Award variables (degree completion and timing). These define the highest degree student i ever earned as of quarter t. The term ''unified'' means highest degree at any school, not just our five institutions. Students who will go on to have a highest degree as a BA are dropped. Each of these are dummies with the omitted group never earning a degree.

   - hdc_short_cert_unified: Highest degree as of quarter t is a short certificate (fewer than 15 credits).
   - hdc_cert_unified: Highest degree as of quarter t is a certificate (more than 15 credits).
   - hdc_assoc_unified: Highest degree as of quarter t is an Associates degree.

3. X_i vector, which are time-invariant characteristics:

   - i.school_age_cohort: Full set of interactions between which of our 5 colleges i enrolled in, dummies for cohort of first enrollment in our colleges (semester), and dummies for age (in years) at first enrollment.
   - i.race_male: Full set of interactions between gender and race (white/non-white).
   - i.pell_loan: Full set of interactions between dummies for ever receiving a Pell grant, or ever taking a loan while enrolled in our schools.
   - non_remedial_cratt_first, remedial_cratt_first: Linear term for number of remedial and non-remedial credits taken (attempted) in first term enrolled.
   - mathZ, englishZ, has_score: First two are Z-scores for test used for english or math remediation. has_score is a dummy for whether i has test scores.

4. W_{it} vector, which are time-varying characteristics:

   - age age2: Age and squared term in each quarter.
   - enrolled_cc, enrolled_nsc: Dummies for enrolled in one of our 5 colleges (enrolled_cc) or enrolled in an NSC school, not including our 5 (enrolled_nsc).
   - hdc_*_enr_cc, hdc_*_enr_nsc : Interactions between time-invariant indicators for highest degree i ever completed and time-varying indicators for whether i is enrolled in one of our 5 colleges (_cc) or an NSC college (_nsc).
   - enr_m1 enr_m2 enr_m3 enr_m4: These are dummies equal to 1 for the 4 quarters prior to enrollment.

5. $X_i t$ vector, which are interactions between time invariant characteristics and a linear time trend.

   - c.trend: Linear term in quarters.

- This is interacted with all elements of the X_i vector.

6. post, post_trend, post_trend x degree.
   - post: A dummy equal to 1 in all quarters after receiving highest degree (for degree earners) or in all quarters after last enrollment for non-completers.
   - post_trend: Defined as the reciprocal of quarters since exit.
   - pt_hdc_short_cert_unified_v2, pt_hdc_cert_unified_v2 pt_hdc_short_cert_unified_v2: are interactions between the post-trend and which degree was earned.
7. α_i is a person specific dummy. In the computation we use Stata's absorb command.
8. τ_t are i.quarter_tq: Secular quarter dummies.
9. $\omega_i\tau$ is a person fixed effect interacted with a linear time trend. We estimate this using Stata's regintfe command as cited in the paper.
10. Matching quarters to semesters. To match quarters to semesters in order to determine enrollment we do the following.
    - If student i was enrolled only in the spring semester (January through May), we defined enrollment in quarter 1 (January-March) equal to 1 (meaning enrolled the entire quarter), and enrollment in quarter 2 (April-June) equal to 2/3 (meaning i was enrolled 2/3rds of the quarter).
    - If i was enrolled only in the summer semester (June-August), we defined enrollment in quarter 2 as 1/3 and enrollment in quarter 3 (July-Sept.) as 2/3.
    - If i was enrolled only in the fall semester (Sept.-Dec.), we defined enrollment in quarter 3 as 1/3, and enrollment in quarter 4 (Oct.-Dec.) as 1.
    - If i was enrolled in consecutive semesters, we would create combinations. For example, enrolling in fall and summer (January-August), enrollment in quarters 1 and 2 would equal 1, and enrollment in quarter 3 would equal 2/3.
11. Higher order trends, and pre- and post- trends. In table 5 we interact our person fixed effect with higher order trend (trend2) and with specific pre- and post- trends.
    - The pre-trend is quarters relative to earning a degree or relative to last quarter enrolled. For example, a pre-trend value of 10 would correspond to the quarter 10 quarters prior to the last quarter i was enrolled.
    - The post-trend is number of quarters since last enrollment or since earning the degree. For example, a value of 5 would correspond to the 5th quarter after earning an award.

0.3 Creating the Plots

The plots are created by generating a measure that is quarters relative to both enrollment and college exit. This variable is negative in quarters until first enrollment (e.g. a value of -10 would correspond to 10 quarters prior to first enrollment) and is positive in quarters since last enrollment or since earning a degree (e.g. a value of 5 would correspond to 5 quarters post degree). This measure takes no value (i.e. is missing) in all quarters after first enrollment and before college exit.

- We then collapse earnings (mean and median) by highest degree ever earned (hdc_ever) and our variable which is quarters relative to enrollment/exit.
- We then plot means by quarters relative to schooling across degree earner categories.

0.4 Summary Statistics Table

This is simply created using the variables described in the regression section. Two additional measures are included which we describe here.

- prior_emp is whether i was employed in the 1 year prior to enrolling.
- prior_earnings_1yr is i's earnings in the year prior to enrollment.

```
* DETAILS OF THE FOLLOWING ARE DESCRIBED IN THE READ-ME FILE ;


*
-------------------------------------------------------------------------------------
*;
* 1 LIMIT SAMPLE;
*
-------------------------------------------------------------------------------------
*;

* DROP IF QUARTER > 20112;
drop if quarter>20112;

* 1ST TIME STUDENTS IN OUR CC's;
keep if ccfirst==1;

* 1ST ENROLLMENT BTW/ 02/07;
keep if first_term_enrolled>=20027 & first_term_enrolled<=20077;

* DROP IF AGE OR GENDER IS MISSING;
drop if age==.;
drop if male==-9;

* KEEP IF 1ST ENROLLED BETWEEN 17 AND45;
keep if age_first_att_credit<=45;
keep if age_first_att_credit>=21;

* AGE SAMPLE RESTRICTION;
keep if age>=17 & age<=65;

* DROP IF EVER ATTENDED/GRADUATED FROM 4-YEAR COLLEGE ;
keep if ever_4year==0 & hdc_ever<4 ;

* DROP IF ENROLLED ON OR AFTER Q1, 2011 ;
keep if last_quarter_attempt_credit_tq<204 ;


* -------------------------------------------------------------------------------------
*    ;
* 2 MAIN REGRESSION TABLES ;
* -------------------------------------------------------------------------------------
*    ;



* GLOBALS ;;
* -------------------------------------------------------------------------------------
*    ;
global awards                "hdc_short_cert_unified hdc_cert_unified
hdc_assoc_unified" ;
global Q                     "i.quarter_tq" ;
global X                     "age age2 enrolled_cc enrolled_nsc i.school_age_cohort
i.race_male i.pell_loan cred_att_first_term c.mathZ c.englishZ has_score enr_m1 enr_m2
enr_m3 enr_m4 hdc*_enr_cc hdc_*_enr_nsc i.school_age_cohort" ;
global Xt                                "enrolled_cc enrolled_nsc  enr_m1 enr_m2 enr_m3
enr_m4 c.trend##(i.race_male i.pell_loan c.non_remedial_cratt_first
c.remedial_cratt_first c.mathZ c.englishZ i.has_score) hdc_*_enr_cc hdc_*_enr_nsc
i.school_age_cohort" ;
global Xt_fe                 "enrolled_cc enrolled_nsc  enr_m1 enr_m2 enr_m3 enr_m4
c.trend#(i.race_male i.pell_loan c.non_remedial_cratt_first c.remedial_cratt_first
c.mathZ c.englishZ i.has_score) hdc_*_enr_cc hdc_*_enr_nsc" ;
global post_deg_v2  "post post_trend_v2 pt_hdc_short_cert_unified_v2
pt_hdc_cert_unified_v2 pt_hdc_short_cert_unified_v2"  ;
```

```
* ------------------------------------------------------------------------------------
*    ;
* 2.1 TABLE 3 ;

eststo clear ;
set more off;

* COL 1 ;
reg ui_quarterly_earnings_real $awards $Q,  vce(cluster id_match)  ;
est sto b1a ;

* COL 2 ;
reg ui_quarterly_earnings_real $awards $Q $X,  vce(cluster id_match)  ;
est sto b2a ;

* COL 3 ;
reg ui_quarterly_earnings_real $awards $Q $Xt,  vce(cluster id_match)  ;
est sto b3a ;

* COL 4 ;
reg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt,   vce(cluster id_match) ;
est sto b4a ;

* COL 5 ;
areg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt_fe,  vce(cluster id_match)
absorb(id_match) ;
est sto b5a ;

* COL 6 ;
regintfe ui_quarterly_earnings_real $awards $post_deg_v2 enrolled_cc enrolled_nsc
hdc_*_enr_cc hdc_*_enr_nsc enr_m* Q_*,  id1(ID) intvar(trend) cluster(ID)  ;
est sto b6a ;

* MAKE TABLE 3 ;
esttab b1a b2a b3a b4a b5a b6a  , se keep(hdc_short_cert_unified hdc_cert_unified
hdc_assoc_unified $post_deg_v2)


* ------------------------------------------------------------------------------------
*    ;
* 2.2 TABLE 4 - BY GENDER ;


* FEMALE ;

* COL 1 ;
reg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt if male==0,   vce(cluster
id_match) ;
est sto b4a_0 ;

* COL 2 ;
areg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt_fe if male==0,
vce(cluster id_match)  absorb(id_match) ;
est sto b5a_0 ;

* COL 3 ;
regintfe ui_quarterly_earnings_real $awards $post_deg_v2 enrolled_cc enrolled_nsc
hdc_*_enr_cc hdc_*_enr_nsc enr_m* Q_* if male==0, id1(ID) intvar(trend) cluster(ID)  ;
est sto b6a_0 ;

* MALE ;

* COL 1 ;
reg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt if male==1,   vce(cluster
id_match) ;
```

```
est sto b4a_1 ;

* COL 2 ;
areg ui_quarterly_earnings_real $awards $post_deg_v2 $Q $Xt_fe if male==1,
vce(cluster id_match)  absorb(id_match) ;
est sto b5a_1 ;

* COL 3 ;
regintfe ui_quarterly_earnings_real $awards $post_deg_v2 enrolled_cc enrolled_nsc
hdc_*_enr_cc hdc_*_enr_nsc enr_m* Q_* if male==1, id1(ID) intvar(trend) cluster(ID)  ;
est sto b6a_1 ;


* MAKE TABLE 4 ;
esttab b4a_0 b5a_0 b6a_0 , se keep(hdc_short_cert_unified hdc_cert_unified
hdc_assoc_unified $post_deg_v2)  ;
esttab b4a_1 b5a_1 b6a_1,  se keep(hdc_short_cert_unified hdc_cert_unified
hdc_assoc_unified $post_deg_v2)  ;



* ------------------------------------------------------------------------------
*    ;
* 2.3 TABLE 5 ;

* COL 1 ;
TAKEN FROM ABOVE ;

* COL 2 ;
regintfe ui_quarterly_earnings_real $awards $post_deg_v2 enrolled_cc enrolled_nsc
hdc_*_enr_cc hdc_*_enr_nsc enr_m* Q_*, id1(ID) intvar(trend trend2) cluster(ID)  ;
est sto b7a ;

* COL 3 ;
regintfe ui_quarterly_earnings_real $awards $post_deg_v2 enrolled_cc enrolled_nsc
hdc_*_enr_cc hdc_*_enr_nsc enr_m* Q_*, id1(ID) intvar(trend_pre trend_post)
cluster(ID)  ;
est sto b8a ;

* MAKE TABLE 5 ;
esttab  b6a b7a b8a , se keep(hdc_short_cert_unified hdc_cert_unified
hdc_assoc_unified $post_deg_v2)  ;



* ------------------------------------------------------------------------------
*    ;
*  3. PLOTS ;
* ------------------------------------------------------------------------------
*    ;


* BELOW WE CREATE A VARIABLE (TMP) WHICH IS ;
* NEGATIVE IN QUARTERS RELATIVE TO ENROLLMENT (E.G. -3 IS 3 QUARTERS UNTIL 1ST
ENROLLMENT);
* POSITIVE IN QUARTERS AFTER LAST ENROLLMENT (E.G. 4 IS 4 QUARTERS AFTER LAST
ENROLLED);
* AND TAKES NO VALUE (=.) IN ALL QUARTERS BETWEEN 1ST AND LAST QUARTERS ENROLLED ;

* DEFINE TMP AS QUARTERS RELATIVE TO FIRST ENROLLMENT (IN NEGATIVE VALUES) ;
cap drop tmp ;
bys id_match: egen tmp=min(cond(enrolled_unified==1, quarter_tq, .)) ;
replace tmp=quarter_tq-tmp ;
replace tmp=. if tmp>=0 ;
```

```
* DEFINE TMP2 AS QUARTERS SINCE LAST ENROLLMENT ;
* REPLACE TMP WITH THIS ;
cap drop tmp2 ;
gen tmp2=relative_degree_or_exit if relative_degree_or_exit<0   ;
replace tmp2=tmp2*-1 ;
replace tmp=tmp2 if tmp==. & tmp2!=. ;
drop tmp2 ;

* KEEP ONLY 3 YEARS PRIOR AND 5 YEARS POST ENROLLMENT ;
replace tmp=. if tmp<-12 ;
replace tmp=. if tmp>20 ;


* MAKE PLOTS BY COLLAPSING OVER QUARTERS UNTIL/AFTER ENROLLMENT;
* DO TWICE, ONCE FOR MEAN, AND ONCE FOR MEDIAN ;
* HDC_EVER IS HIGHEST DEGREE EVER ;

preserve ;
collapse ui_quarterly_earnings_real, by(tmp hdc_ever) ;

twoway
(connected ui_quarterly_earnings_real tmp if hdc_ever==0 & tmp<0, lc(gs3) mc(gs3)
ms(o) lp(solid))
(connected ui_quarterly_earnings_real tmp if hdc_ever==1 & tmp<0, lc(gs6) mc(gs6)
ms(d) lp(longdash dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==2 & tmp<0, lc(gs8) mc(gs8)
ms(s) lp(longdash))
(connected ui_quarterly_earnings_real tmp if hdc_ever==3 & tmp<0, lc(gs10) mc(gs10)
ms(t) lp(dash_dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==0 & tmp>0, lc(gs3) mc(gs3)
ms(o) lp(solid))
(connected ui_quarterly_earnings_real tmp if hdc_ever==1 & tmp>0, lc(gs6) mc(gs6)
ms(d) lp(longdash dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==2 & tmp>0, lc(gs8) mc(gs8)
ms(s) lp(longdash))
(connected ui_quarterly_earnings_real tmp if hdc_ever==3 & tmp>0, lc(gs10) mc(gs10)
ms(t) lp(dash_dot))
, ysca(r(0 8000))  ylab(0(2000)8000) scheme(s1mono) xline(-1, lp(dot)) xline(1,
lp(dot))
xlabel(-10 "-10" -5 "-5" -1 "First" 1 "Last" 5 "5" 10 "10" 15 "15" 20 "20",
labsize(small) angle(forty_five))
legend(order(1 "No degree" 2 "Short Cert." 3 "Certificate" 4 "Associate's") col(3))
ytitle("Average (real) quarterly earnings" " ") xtitle(Quarters relative to
enrollment/exit) ysize(3) xsize(3);
restore ;


preserve ;
collapse (median) ui_quarterly_earnings_real, by(tmp hdc_ever) ;
twoway
(connected ui_quarterly_earnings_real tmp if hdc_ever==0 & tmp<0, lc(gs3) mc(gs3)
ms(o) lp(solid))
(connected ui_quarterly_earnings_real tmp if hdc_ever==1 & tmp<0, lc(gs6) mc(gs6)
ms(d) lp(longdash dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==2 & tmp<0, lc(gs8) mc(gs8)
ms(s) lp(longdash))
(connected ui_quarterly_earnings_real tmp if hdc_ever==3 & tmp<0, lc(gs10) mc(gs10)
ms(t) lp(dash_dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==0 & tmp>0, lc(gs3) mc(gs3)
ms(o) lp(solid))
(connected ui_quarterly_earnings_real tmp if hdc_ever==1 & tmp>0, lc(gs6) mc(gs6)
ms(d) lp(longdash dot))
(connected ui_quarterly_earnings_real tmp if hdc_ever==2 & tmp>0, lc(gs8) mc(gs8)
ms(s) lp(longdash))
(connected ui_quarterly_earnings_real tmp if hdc_ever==3 & tmp>0, lc(gs10) mc(gs10)
ms(t) lp(dash_dot))
```

```
, ysca(r(0 8000))  ylab(0(2000)8000) scheme(s1mono) xline(-1, lp(dot)) xline(1,
lp(dot))
xlabel(-10 "-10" -5 "-5" -1 "First" 1 "Last" 5 "5" 10 "10" 15 "15" 20 "20",
labsize(small) angle(forty_five))
legend(order(1 "No degree" 2 "Short Cert." 3 "Certificate" 4 "Associate's") col(3))
ytitle("Median (real) quarterly earnings" " ") xtitle(Quarters relative to enrollment/
exit)  ysize(3) xsize(3); ;
restore ;
```

```
* ------------------------------------------------------------------------------------
*    ;
*  4. SUMMARY STAT TABLE (TABLE 2) ;
* ------------------------------------------------------------------------------------
*    ;
cap restore, not ;
preserve  ;

* DON'T INCLUDE 0's FOR MISSING (DUMMIED OUT) MATH/ENGLISH SCORES ;
replace englishZ=. if english_miss==1 ;
replace mathZ=. if math_miss==1 ;

* KEEP ONE OBS PER PERSON (ALL MEASURES ARE TIME-INVARIANT);
bys id_match: keep if _n==1 ;

forvalues i=0(1)3 { ;
          tabstat age_first_att_credit white non_white race_miss ever_pell ever_loan
          non_remedial_cratt_first remedial_cratt_first prior_emp prior_earnings_1yr
          englishZ mathZ english_miss math_miss if hdc_ever==`i', stat(mean) save   ;
          tabstatmat col`i' ;
          mat col`i'=col`i'' ;
} ;
cap drop tmp_* ;
gen tmp=1 ;

tabstat tmp, by(hdc_ever) stat(N) notot save ;
tabstatmat count ;
mat table=(col0, col1, col2, col3 \ count') ;
mat list table ;
```