Appendix

The Effects of High School Career and Technical Education on Employment, Wages, and Educational Attainment

Michael LaForest
Pennsylvania State University
mlaforest@psu.edu

February 25, 2022

Table of Contents

A. Summary Statistics – Education and Labor Market Pathways

A.1 Educational Attainment and Wages

Table A1: High School Curricula (By Age 19)

HS Curriculum	# Obs	% Sample
Academic	3,370	21%
General Education	5,320	33%
Business Vocational	870	5%
Trade Vocational	720	5%
Other	2,130	13%
HS Graduate, Curr Unknown	1,960	12%
GED	470	3%
No HS Degree	1,070	7%

Notes:

1) Total # observations is 15,900. HS graduation is unknown for remaining 300 sample members.

2) Table refers to HS graduation attainment through the '04-'05 academic year, 5 years after sample members began high school.

3) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

Table A2: Educational Attainment (By Age 26)

PSE Attainment	# Obs	% Sample	2012 CPS 25-34 Yr Olds
No HS Graduation	450	3%	11%
HS Graduation Only	5,420	40%	45%
1-yr Trade School	1,230	9%	--
2-yr Community College	1,050	8%	10%
4-yr University	5,100	38%	34%

Notes:

1) Total # observations is 13,250. Educational attainment is unknown for 2,340 sample members.

2) The CPS does not collect vocational certificate information.

3) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

Table A3: Log Hourly Wages

Variable	Mean	Std Dev	# Obs
(ln) Professional Hr Wage	2.666	0.493	6,310
(ln) Skilled Manual Labor Hr Wage	2.430	0.437	5,870
(ln) Skilled Non-Manual Labor Hr Wage	2.299	0.414	7,120
(ln) Skilled Other Hr Wage	2.502	0.460	1,140
(ln) Unskilled Hr Wage	2.073	0.357	2,290

Notes:
1) Each observation is an individual-year log hourly wage. Log hourly wages are constructed by first converting all wages that were recorded over the length of the survey into real 2002 dollars. Wages are then converted into hourly wages and logged.
2) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

A.2 Education and Labor Market Pathways

Table A4 breaks down PSE degree attainment by high school curriculum. First, 73% of individuals who took academic courses completed four-year university degrees. Comparatively, 59% of individuals who took general education courses, 41% of individuals who took trade vocational courses, 51% of individuals who took business vocational courses, and 7% of individuals who had not graduated high school by age 19 had graduated from at least one type of PSE institution by the time the study concluded in 2012.

Table A5 displays 2012 employment outcomes by high school curriculum. Overall, academic concentrators were most likely to later work in professional occupations (49%), trade vocational concentrators were most likely to later work in skilled manual labor occupations (45%), and business vocational concentrators were most likely to later work in skilled non-manual labor occupations (35%).

Finally, Table A6 includes the aggregate percentage breakdown of individual choices between 2000 and 2012. The majority of individuals attended high school between 2000 and 2003, and those who attended PSE institutions mostly did so between 2004 and 2008. Note that ELS:2002 asked very few labor market questions about the period between 2006-2010. While some of these values are imputed based on job start and end dates, many of them are coded as missing or "Work Unknown Type" during these years.[1]

[1] Additional details about imputation rules are provided in Appendix F. In addition to the observed choices described in Table A6, I observe information about whether some individuals never graduate from high school, never attain a GED, or never graduate from a particular kind of PSE institution. This information is used when calculating the likelihood functions of individuals with missing information as described in Appendix E.

Table A4: PSE Attainment (Age 26) by HS Curriculum

HS Curriculum	No PSE *% Sample*	1-yr Trade School *% Sample*	2-yr Community College *% Sample*	4-yr University *% Sample*	# Obs
Academic	19%	4%	4%	73%	3,050
Gen Ed	41%	10%	10%	38%	4,310
Bus Voc	49%	10%	11%	30%	710
Trade Voc	59%	13%	12%	17%	570
Other	53%	14%	10%	22%	1,690
GED (By Age 19)	74%	15%	5%	5%	390
No HS Degree (By Age 19)	93%	5%	2%	0%	870

Notes:

1) Percentages aggregate left to right.

2) Total # observations is 12,580. HS Curriculum and/or PSE attainment is unknown for 3,620 sample members.

3) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

Table A5: Employment Outcomes (Age 26) by HS Curriculum

HS Curriculum	Professional *% Sample*	Sk. Manual Labor *% Sample*	Sk. Non-Manual Labor *% Sample*	Sk. Other *% Sample*	Unskilled *% Sample*	Not Employed *% Sample*	# Obs
Academic	49%	13%	21%	7%	4%	5%	2,690
Gen Ed	29%	20%	30%	5%	7%	9%	3,880
Bus Voc	27%	19%	35%	2%	7%	9%	650
Trade Voc	18%	45%	20%	3%	7%	8%	520
Other	20%	23%	30%	5%	9%	12%	1,500
GED (By Age 19)	18%	24%	28%	2%	11%	16%	340
No HS Degree (By Age 19)	11%	28%	22%	1%	13%	26%	760

Notes:

1) Percentages aggregate left to right.

2) Total # observations is 10,330. HS Curriculum and/or 2012 employment is unknown for 5,870 sample members.

3) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

Table A6: ELS:2002 Choices By Year

Choices by year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
HS Academic	760	1,820	2,920	3,140	-	-	-	-	-	-	-	-	-
HS Gen Ed	12,100	10,830	8,330	5,520	-	-	-	-	-	-	-	-	-
HS Business Voc	220	280	520	810	-	-	-	-	-	-	-	-	-
HS Trade Voc	200	390	640	630	-	-	-	-	-	-	-	-	-
HS Other	800	850	1,070	2,010	-	-	-	-	-	-	-	-	-
GED	-	10	100	230	140	90	50	50	30	40	30	40	10
HS UNKNOWN TYPE	1,660	1,510	1,740	1,610	330	80	30	10	-	-	-	-	-
WORK Professional	-	-	-	30	210	320	500	300	730	1,160	1,690	2,480	3,550
WORK Skilled Manual Labor	-	10	40	110	720	960	1,380	610	820	1,010	1,300	1,720	2,440
WORK Skilled Non-Manual Labor	-	20	30	90	650	1020	1,470	640	930	1,270	1,670	2,250	3,220
WORK Skilled Other	-	-	-	10	90	100	130	60	160	270	360	460	570
WORK Unskilled	-	70	60	150	590	630	820	220	270	330	410	510	820
WORK UNKNOWN TYPE	80	280	60	520	400	440	10	10	10	4,000	3,260	1,730	300
UNEMPLOYED	-	-	60	590	780	450	550	150	220	1,170	1,040	1,010	1,200
PSE 1YR Trade School	-	-	-	20	170	210	170	190	180	150	160	100	70
PSE 2YR Community College	-	-	10	70	2,810	1,650	500	390	370	380	370	290	110
PSE 4YR University	-	-	-	60	6,730	6,440	4,370	3,280	2,310	1,250	950	780	220
PSE UNKNOWN TYPE	-	-	-	-	10	70	60	50	80	50	40	30	40
PSE HIGHER DEGREE	-	-	-	-	-	10	-	10	110	120	160	130	320
MISSING	380	130	630	600	2,580	3,740	6,140	10,230	9,990	5,010	4,770	4,660	3,340

Notes:

1) Total # Observations is 16200.

2) Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

B. Two Stage Least Squares Regression Results

B.1 First Stage Results - High School Curriculum

The first-stage regression, used to construct high school curriculum predicted probabilities, is a multinomial logit regression of high school curriculum on personal characteristics (X_i), local labor market characteristics (M_i), and high school vocational instruments (I_i). The estimates from the first-stage regression are displayed in Table B1; note that all estimates are relative to graduating high school in a general education curriculum.[2] Overall, men are more likely to concentrate in a trade vocational field then women. Specifically, men receive 1.43 more utils than women from concentrating in the trade vocational curriculum relative to concentrating in the general education curriculum. Next, Caucasian individuals are more likely than black, Hispanic, and other race individuals to concentrate in a trade vocational, business vocational, or other curriculum, and individuals who attend Catholic or non-Catholic private high schools are very likely to take general education courses as opposed to academic or vocational courses and are also very unlikely to drop out of high school. Local labor market characteristics have little effect on curriculum take-up, although there are a few exceptions. For example, as the hourly wage increases in the county where the school is located, the number of individuals who concentrate in other curricula decreases, and as the percent of manual labor employment increases in the county where the school is located, the number of individuals who concentrate in a trade vocational curriculum increases.

Each instrument has a significant effect on the utility associated with at least one high school curriculum relative to graduating in the general education field, with the exception of whether most vocational courses are taught in the high school, at an area vocational school, or both (this variable has a positive but statistically insignificant effect on concentrating in a trade or business vocational curriculum).[3] As the number of individuals in the previous year's graduating class who took vocational courses increases, the probability that an individual concentrates in a trade vocational curriculum or a business vocational curriculum increases. Next, when business courses such as marketing are taught on-site, individuals are more likely to concentrate in business vocational fields. When trade courses such as precisions are taught on-site, individuals are more

[2] Note that the results, described below, are appreciably similar when estimated using a control function approach.

[3] The first-stage estimates for vocational course location are significant for many alternative specifications of the instrument subset. Changing the instrument subset has little effect on the estimates in the second-stage regressions. The instrument subset presented here was chosen to be indicative of the full set of variables available in ELS:2002.

Table B1: Selected First Stage Estimates

Variable	Academic Estimate	Academic SE	Business Voc Estimate	Business Voc SE	Trade Voc Estimate	Trade Voc SE	Other Estimate	Other SE	GED Estimate	GED SE	Dropping Out Estimate	Dropping Out SE
1. Personal Characteristics												
Male	**-0.36** ***	*0.06*	-0.05	*0.08*	**1.43** ***	*0.11*	**-0.18** ***	*0.05*	**0.62** ***	*0.11*	**0.42** ***	*0.08*
Black	0.02	*0.12*	-0.15	*0.15*	**-0.83** ***	*0.17*	**-0.37** ***	*0.10*	**-0.62** ***	*0.19*	-0.22	*0.15*
Hispanic	-0.14	*0.10*	**-0.56** ***	*0.18*	**-0.83** ***	*0.18*	**-0.36** ***	*0.10*	-0.23	*0.20*	0.21	*0.13*
Other Race	**0.60** ***	*0.08*	-0.06	*0.14*	**-0.39** ***	*0.15*	**-0.36** ***	*0.11*	-0.04	*0.18*	0.19	*0.14*
Socio-Economic Status	**0.24** ***	*0.03*	**-0.11** **	*0.05*	**-0.16** ***	*0.06*	-0.04	*0.03*	**-0.18** ***	*0.06*	**-0.39** ***	*0.05*
Test Score	**1.34** ***	*0.04*	-0.07	*0.05*	**-0.36** ***	*0.05*	**-0.29** ***	*0.04*	**-0.14** **	*0.07*	**-0.66** ***	*0.05*
Midwest	**-0.43** ***	*0.13*	-0.22	*0.19*	**-0.41** **	*0.19*	**0.27** *	*0.15*	-0.24	*0.26*	**-0.43** **	*0.19*
South	0.08	*0.13*	-0.03	*0.18*	**-0.56** ***	*0.20*	0.03	*0.15*	0.16	*0.25*	**-0.39** **	*0.19*
West	**-0.65** ***	*0.14*	**-1.29** ***	*0.25*	**-0.85** ***	*0.26*	-0.03	*0.19*	**-0.63** **	*0.28*	**-0.84** ***	*0.23*
Suburban	**-0.33** ***	*0.10*	-0.06	*0.15*	0.24	*0.16*	0.11	*0.11*	-0.12	*0.16*	-0.20	*0.13*
Rural	**-0.46** ***	*0.14*	-0.06	*0.18*	-0.10	*0.21*	0.12	*0.14*	-0.16	*0.24*	-0.01	*0.20*
Catholic School	**-0.83** ***	*0.19*	0.18	*0.29*	**-0.85** **	*0.36*	**-1.44** ***	*0.27*	**-1.33** ***	*0.41*	**-1.74** ***	*0.32*
Non-Catholic Private School	**-1.07** ***	*0.21*	**-0.78** **	*0.32*	**-1.32** ***	*0.41*	**-1.46** ***	*0.27*	**-0.83** **	*0.39*	**-1.01** ***	*0.28*
2. Local Labor Market Characteristics												
Unemployment Rate	-0.60	*3.12*	4.83	*4.06*	-1.57	*5.04*	-1.01	*2.80*	-0.83	*4.26*	-3.58	*3.01*
(ln) Average Hourly Wage	0.08	*0.28*	-0.08	*0.41*	-0.27	*0.45*	**-0.74** **	*0.32*	-0.39	*0.49*	-0.32	*0.37*
% Professional Employment	0.40	*2.09*	0.81	*2.73*	-0.65	*3.59*	0.89	*2.31*	-5.16	*3.54*	-3.11	*2.99*
% Manual Labor Employment	-0.32	*0.70*	0.54	*0.90*	**2.25** **	*1.10*	**-1.27** *	*0.67*	-1.21	*0.99*	-1.10	*0.93*
% Non-Manual Labor Employment	-0.10	*1.22*	1.87	*1.74*	0.17	*2.05*	-1.31	*1.21*	-1.43	*1.85*	0.64	*1.63*
3. Vocational Instruments												
Voc Taught in High School	-0.05	*0.16*	0.37	*0.31*	0.43	*0.30*	0.13	*0.16*	-0.24	*0.36*	0.26	*0.25*
Voc Taught in Area School	**-0.29** **	*0.14*	0.21	*0.35*	0.16	*0.28*	0.03	*0.17*	-0.29	*0.43*	**0.45** *	*0.26*
Voc Taught in Both HS & Area Sch	-0.14	*0.16*	0.33	*0.32*	0.22	*0.30*	0.14	*0.17*	-0.15	*0.38*	0.30	*0.25*
Marketing Courses Taught On-Site	-0.03	*0.12*	**0.61** ***	*0.17*	0.15	*0.17*	**0.20** *	*0.11*	**0.36** **	*0.18*	0.06	*0.14*
Marketing Courses Taught at Area Sch	**-0.31** *	*0.18*	0.08	*0.25*	0.04	*0.27*	0.09	*0.16*	-0.01	*0.28*	-0.24	*0.24*
Precisions Courses Taught On-Site	0.15	*0.14*	-0.16	*0.20*	**0.42** **	*0.20*	-0.06	*0.14*	-0.02	*0.23*	0.18	*0.19*
Precisions Courses Taught at Area Sch	-0.03	*0.16*	-0.13	*0.22*	**0.41** *	*0.24*	-0.14	*0.17*	-0.03	*0.26*	0.16	*0.21*
# Vocational Teachers per 100 Students	**-0.26** **	*0.11*	0.07	*0.13*	**0.16** *	*0.09*	0.06	*0.08*	-0.10	*0.13*	0.03	*0.11*
Career Pathways Prog Available	0.14	*0.10*	**0.31** *	*0.17*	0.26	*0.19*	-0.02	*0.11*	-0.09	*0.45*	0.03	*0.21*
% Students Free/Reduced Price Lunch	0.12	*0.22*	-0.10	*0.32*	0.31	*0.31*	-0.07	*0.21*	0.08	*0.35*	**0.46** *	*0.25*
% Students Take Academic Courses	-0.06	*0.20*	**-0.55** **	*0.27*	-0.20	*0.31*	**-0.62** ***	*0.17*	**-1.88** ***	*0.70*	0.16	*0.35*
% Students Take Vocational Courses	0.45	*0.45*	**1.00** ***	*0.37*	**1.88** ***	*0.43*	**1.19** ***	*0.34*	**1.92** **	*0.90*	0.42	*0.53*
% Prev Students Enter Labor Market (0-5)	-0.04	*0.07*	0.12	*0.09*	0.10	*0.10*	0.06	*0.06*	**-0.41** *	*0.24*	**-0.22** **	*0.11*
Admission Based on Geography	0.04	*0.13*	0.06	*0.19*	0.13	*0.23*	-0.04	*0.13*	0.60	*0.59*	**-0.53** **	*0.24*
Student Infl on Course Selection (0-3)	0.01	*0.06*	**0.17** *	*0.10*	0.15	*0.11*	0.08	*0.07*	-0.25	*0.21*	0.10	*0.15*
GED Confered by High School	-0.03	*0.14*	-0.12	*0.17*	-0.08	*0.19*	**-0.34** ***	*0.11*	**0.90** **	*0.35*	**0.48** **	*0.21*
Constant	-0.20	*0.91*	**-2.87** **	*1.43*	**-4.15** ***	*1.39*	**1.67** *	*1.00*	-2.12	*2.07*	-1.23	*1.29*

Notes:

1) Multinomial Logit regression. Estimates are relative to graduating high school in the general education field.

2)*,**,*** denote 90%, 95%, and 99% statistical significance respectively.

3) Standard Errors (SE) are clustered at the school level.

4) Total # Observations is 15,890.

likely to concentrate in trade vocational fields. Offering career pathways programs to students increases their likelihood of taking business vocational courses, while increasing the number of vocational teachers per student at a school increases the likelihood of taking trade vocational courses. As well, students who attend schools that confer GEDs on-site are more likely to pursue GEDs and are also slightly more likely to drop out of high school. Whether schools admit students based on geographic location has little effect on curriculum choice, with the exception that students are less likely to pursue GEDs or dropout. Finally, as students' influence on course selection increases, students are more likely to take business vocational courses relative to general education courses and are less likely to pursue GEDs.

B.2 First Stage Results - PSE Attainment

Table B2 provides the results of a first first-stage regression used to construct post-secondary education attainment predicted probabilities. The regression is a multinomial logit regression of PSE attainment on personal characteristics, local labor market characteristics, and high school instruments related to PSE attendance and PSE opportunities.

Table B2: Selected PSE First Stage Estimates

	1-yr Trade School			2-yr CC			4-yr University		
Variable	Estimate		*SE*	Estimate		*SE*	Estimate		*SE*
1. Personal Characteristics									
Male	**-0.63**	***	*0.06*	**-0.47**	***	*0.07*	**-0.60**	***	*0.04*
Black	0.07		*0.10*	**-0.31**	***	*0.12*	0.05		*0.09*
Hispanic	**-0.21**	**	*0.10*	-0.14		*0.11*	**-0.16**	*	*0.09*
Other Race	-0.10		*0.11*	**-0.21**	*	*0.11*	**0.37**	***	*0.07*
Socio-Economic Status	0.04		*0.04*	**0.12**	***	*0.04*	**0.51**	***	*0.03*
Test Score	**-0.08**	**	*0.04*	**0.15**	***	*0.04*	**0.94**	***	*0.03*
Midwest	0.04		*0.10*	-0.01		*0.11*	**-0.26**	***	*0.08*
South	0.04		*0.10*	**-0.19**	*	*0.11*	**-0.24**	***	*0.08*
West	0.10		*0.11*	-0.16		*0.14*	**-0.44**	***	*0.09*
Suburban	-0.02		*0.08*	0.04		*0.09*	**-0.17**	***	*0.06*
Rural	-0.02		*0.11*	0.11		*0.13*	**-0.18**	**	*0.09*
Catholic School	**0.31**	**	*0.15*	**0.29**	*	*0.18*	**0.44**	***	*0.11*
Non-Catholic Private School	0.18		*0.15*	**0.42**	**	*0.17*	**0.41**	***	*0.11*
2. Local Labor Market Characteristics									
Unemployment Rate	2.56		*2.20*	0.43		*2.58*	**7.44**	***	*2.10*
(ln) Average Hourly Wage	0.40		*0.21*	0.12		*0.28*	0.10		*0.19*
% Professional Employment	-1.95		*1.86*	0.86		*1.95*	**3.63**	**	*1.46*
% Manual Labor Employment	1.34		*0.48*	0.70		*0.59*	0.56		*0.44*
% Non-Manual Labor Employment	0.02		*0.89*	0.86		*1.12*	0.35		*0.75*
3. High School PSE Instruments									
% Students Attend College Fairs	-0.03		*0.03*	0.00		*0.04*	0.03		*0.02*
% Students in College App Prog (0-5 Scale)	0.01		*0.03*	0.01		*0.03*	-0.02		*0.02*
% Prev Students Enter Labor Market (0-5 Scale)	0.00		*0.05*	-0.03		*0.06*	-0.05		*0.04*
% Prev Students Attend 2yr College (0-5 Scale)	**0.13**	***	*0.05*	**0.16**	***	*0.05*	0.03		*0.03*
% Prev Students Attend 4yr College (0-5 Scale)	0.02		*0.04*	-0.09		*0.06*	**0.12**	***	*0.04*
% Students Free/Reduced Price Lunch	-0.03		0.15	-0.09		0.18	**-0.29**	*	0.16
% Students Take Academic Courses	-0.13		*0.16*	0.07		*0.20*	**-0.28**	**	*0.14*
Admission Based on Geography	0.03		*0.13*	0.07		*0.13*	0.10		*0.08*
Student Infl on Course Selection (0-3)	-0.01		*0.06*	**0.11**	*	*0.06*	0.07		*0.04*
Constant	**-2.80**	***	*0.68*	**-2.18**	**	*0.91*	**-1.14**	***	*0.63*

Notes:
1) Multinomial Logit regression. Estimates are relative to no PSE attainment.
2)*,**,*** denote 90%, 95%, and 99% statistical significance respectively.
3) Standard Errors (SE) are clustered at the school level.
4) Total # Observations is 13,250.

B.3 Second Stage Estimates

Table B3 presents the 2SLS estimates. Column 1 presents estimates from an OLS regression of log hourly wages on high school curriculum, without instruments, that does not account for curricula self-selection or post-secondary education attainment. Column 2 presents 2SLS estimates from a second-stage OLS regression of log hourly wages on high school predicted probabilities and PSE predicted probabilities from two separate first stage regressions.[4]

Columns 3 and 4 present results from two different second-stage logit regressions – whether or not an individual is employed at age 26 (Column 3), and whether or not an individual is employed in a skilled occupation at age 26 conditional on employment (Column 4), on high school curricula and PSE predicted probabilities.[5]

[4] Note that each result is robust to choosing different subsets of instruments in the first-stage regression, with the exception of the estimate for business vocational curricula (which is always negative but whose statistical significance varies across regressions as I choose different subsets of instruments).
[5] Note that the results in Columns 3 and 4 are robust to choosing different subsets of instruments in the first-stage regression with two exceptions: the skilled occupation parameter estimates for community college and four-year university graduation vary in significance as I run the regressions on different subsets of instruments (though the estimate on community college always has a negative sign and the estimate on four-year university always has a positive sign).

Table B3: HS Curricula on Labor Market Outcomes (Age 26)

	Wages (w/o instruments)			Wages			Employed			Skilled Occupation		
Variable	Estimate		*SE*	Estimate		*SE*	Estimate		*SE*	Estimate		*SE*
Prob Academic	**0.13**	***	*(.01)*	-0.02		*(.07)*	-0.81		*(.55)*	0.12		*(.56)*
Prob Business Vocational	**0.06**	***	*(.02)*	**-0.45**	**	*(.21)*	0.74		*(1.38)*	1.29		*(1.78)*
Prob Trade Vocational	**0.06**	**	*(.02)*	**0.34**	***	*(.12)*	0.77		*(.83)*	**3.03**	***	*(1.04)*
Prob Other Curriculum	-0.01		*(.02)*	0.01		*(.11)*	**-2.55**	***	*(.69)*	-0.45		*(.86)*
Prob GED	**-0.08**	***	*(.02)*	0.12		*(.20)*	-0.40		*(1.28)*	-0.19		*(1.35)*
Prob HS Dropout	**-0.17**	***	*(.02)*	**-0.23**	***	*(.08)*	**-1.92**	***	*(.55)*	-0.52		*(.68)*
Prob 1-yr Trade School	**-**		-	**-0.45**	*	*(.23)*	2.51		*(1.64)*	-0.88		*(1.94)*
Prob 2-yr Community College	**-**		-	0.00		*(.21)*	2.28		*(1.42)*	**-3.57**	**	*(1.74)*
Prob 4-yr University	**-**		-	**0.32**	***	*(.09)*	**2.33**	***	*(.69)*	1.26		*(.81)*
Male	**0.08**	***	*(.01)*	**0.06**	***	*(.01)*	**0.91**	***	*(.12)*	0.18		*(.14)*
Black	**-0.08**	***	*(.01)*	**-0.07**	***	*(.02)*	0.15		*(.12)*	0.15		*(.14)*
Hispanic	-0.01		*(.01)*	-0.01		*(.02)*	**0.24**	**	*(.12)*	0.18		*(.13)*
Other Race	0.02		*(.01)*	0.01		*(.02)*	**-0.18**	*	*(.11)*	-0.18		*(.13)*
Socio-Economic Status	**0.04**	***	*(.01)*	0.01		*(.01)*	-0.01		*(.06)*	-0.08		*(.07)*
Testscore	**0.07**	***	*(.01)*	**0.03**	**	*(.01)*	0.09		*(.09)*	0.10		*(.11)*
Midwest	**-0.07**	***	*(.02)*	**-0.06**	***	*(.02)*	**0.33**	**	*(.13)*	-0.05		*(.15)*
South	**-0.08**	***	*(.02)*	**-0.06**	***	*(.02)*	0.08		*(.11)*	-0.17		*(.14)*
West	-0.03		*(.02)*	-0.02		*(.02)*	-0.03		*(.14)*	-0.01		*(.16)*
Suburban	**0.03**	**	*(.01)*	**0.03**	**	*(.01)*	0.07		*(.08)*	0.03		*(.10)*
Rural	**0.03**	*	*(.02)*	**0.03**	*	*(.02)*	0.09		*(.11)*	0.23		*(.15)*
Catholic School	**0.08**	***	*(.02)*	**0.05**	**	*(.02)*	-0.14		*(.19)*	0.09		*(.20)*
Non-Catholic Private School	**0.07**	***	*(.02)*	0.02		*(.02)*	**-0.81**	***	*(.17)*	**0.52**	**	*(.23)*
Unemployment Rate	0.22		*(.34)*	0.16		*(.36)*	-1.60		*(2.63)*	0.86		*(2.46)*
(ln) Average Hourly Wage	**0.07**	*	*(.04)*	**0.08**	**	*(.04)*	**-0.49**	**	*(.24)*	0.24		*(.28)*
% Professional Employment	**0.61**	**	*(.28)*	0.37		*(.29)*	**3.78**	*	*(2.11)*	-0.10		*(2.04)*
% Manual Labor Employment	0.09		*(.09)*	0.12		*(.10)*	**-1.72**	***	*(.62)*	**1.26**	*	*(.73)*
% Non-Manual Labor Employment	0.02		*(.14)*	0.06		*(.14)*	0.17		*(1.02)*	**2.07**	*	*(1.21)*
Constant	**2.17**	***	*(.11)*	**2.12**	***	*(.12)*	**2.77**	***	*(.83)*	0.65		*(1.01)*

Notes:

1) Column 1 presents Ordinary Least Squares (OLS) regression estimates of log hourly wage on high school curriculum, without instruments. Columns 2-4 present regression estimates of log hourly wages (OLS), employment (logit - employed (1) vs not employed (0)), and skilled employment (logit - employed in skilled occupation (1) vs employed in unskilled occupation (0), conditional on working) on high school predicted probabilities and PSE predicted probabilities from two separate first stage regressions. See Table 4.3 for a list of instruments and appendix Tables E.1 and E.2 for first stage regression results.

2) HS predicted probabilities are relative to graduating high school in the general education field.

3)*,**,*** denote 90%, 95%, and 99% statistical significance respectively.

4) Standard Errors are clustered at the school level.

5) Total # observations is 10,020 for regressions 1 and 2, 12,100 for regression 3, and 10,590 for regression 4.

C. Structural Estimates

Tables C1, C2, and C3, display the structural parameter estimates not provided in Section 6.2. Table C1 displays the omitted parameters related to the five occupation choices in the model. Table C2 displays the omitted parameters related to the three PSE institution choices in the model. Table C3 displays the omitted parameters related to the five high school field choices and the GED choice in the model.

As shown in Table C1, men receive higher wages than women in every occupation except the skilled non-manual labor occupation, and wages tend to increase on average as an individual's socio-economic status and test scores increase. In addition, the non-pecuniary utility of each occupation, relative to choosing not to work, also increases as an individual's socio-economic status and test scores increase.

As shown in Table C3, women receive higher non-pecuniary utility than men in all high school fields except the trade vocational field, and individuals with high socio-economic status and test scores receive a large amount of non-pecuniary utility from attending high school in any field relative to dropping out of high school.

Table C1: Additional Structural Occupation Parameters

	Professional		Skilled Manual Labor		Skilled Non-Manual Labor		Skilled Other		Unskilled	
Variable	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
1. HS Curriculum & 1-yr Trade School Complementarity (Log-Wage Utility)										
Business Voc & 1-yr PSE	-	-	-0.01	*(.028)*	-0.02	*(.025)*	-	-	-	-
Trade Voc & 1-yr PSE	-	-	0.05 *	*(.030)*	0.11 ***	*(.039)*	-	-	-	-
2. Local Labor Market Characteristics (Log-Wage Utility)										
Unemployment Rate	-0.12	*(.274)*	0.66 ***	*(.179)*	-0.62 ***	*(.212)*	1.35 ***	*(.472)*	-1.40 ***	*(.309)*
(ln) Average Hourly Wage	0.10 ***	*(.024)*	0.03 *	*(.017)*	0.13 ***	*(.020)*	-0.07	*(.044)*	0.11 ***	*(.032)*
% Professional Emp	0.90 ***	*(.197)*	0.59 ***	*(.141)*	-0.55 ***	*(.153)*	0.52	*(.342)*	-0.84 ***	*(.252)*
% Manual Labor Emp	0.22 ***	*(.060)*	0.25 ***	*(.044)*	-0.19 ***	*(.049)*	0.07	*(.100)*	-0.63 ***	*(.074)*
% Non-Manual Labor Emp	-0.07	*(.104)*	-0.08	*(.081)*	0.44 ***	*(.087)*	0.44 **	*(.195)*	0.11	*(.138)*
3. Personal Characteristics (Log-Wage Utility)										
Male	0.05 ***	*(.011)*	0.22 ***	*(.013)*	-0.04 ***	*(.011)*	0.09 ***	*(.024)*	0.07 ***	*(.018)*
Black	-0.09 ***	*(.022)*	-0.03	*(.018)*	0.01	*(.018)*	0.16 ***	*(.039)*	-0.07 **	*(.029)*
Hispanic	-0.06 ***	*(.019)*	-0.04 *	*(.018)*	0.04 *	*(.018)*	0.10 ***	*(.039)*	0.01	*(.026)*
Other Race	0.05 ***	*(.016)*	-0.06 ***	*(.018)*	0.05 ***	*(.018)*	0.15 ***	*(.037)*	0.08 ***	*(.027)*
Socio-Economic Status	0.03 ***	*(.006)*	0.00	*(.007)*	-0.01	*(.007)*	0.00	*(.014)*	0.03 **	*(.011)*
Testscore	0.06 ***	*(.007)*	0.03 ***	*(.007)*	0.02 ***	*(.007)*	0.04 ***	*(.014)*	-0.01	*(.011)*
Midwest	-0.10 ***	*(.016)*	-0.04 **	*(.018)*	-0.04 **	*(.017)*	-0.16 ***	*(.035)*	-0.03	*(.029)*
South	-0.09 ***	*(.016)*	-0.05 ***	*(.018)*	-0.02	*(.017)*	-0.10 ***	*(.033)*	-0.05 *	*(.027)*
West	-0.02	*(.017)*	0.04	*(.019)*	0.03	*(.019)*	-0.20 ***	*(.041)*	-0.07 **	*(.030)*
Suburban	0.02	*(.013)*	0.05 ***	*(.014)*	-0.01	*(.013)*	0.02	*(.028)*	0.01	*(.021)*
Rural	0.01	*(.017)*	0.02	*(.018)*	0.02	*(.018)*	-0.06	*(.034)*	-0.07 **	*(.027)*
Catholic School	0.07 ***	*(.016)*	0.11 ***	*(.024)*	0.06 ***	*(.021)*	-0.04	*(.038)*	-0.02	*(.038)*
Non-Catholic Private Sch	0.02	*(.018)*	0.03	*(.024)*	0.09 ***	*(.021)*	0.09 **	*(.037)*	0.07	*(.043)*
4. Personal Characteristics (Non-Pecuniary Utility)										
Male	0.07 ***	*(.020)*	0.91 ***	*(.021)*	-0.08 ***	*(.018)*	-0.18 ***	*(.036)*	-0.15 ***	*(.031)*
Black	0.00	*(.036)*	-0.33 ***	*(.029)*	0.05	*(.029)*	-0.16 ***	*(.061)*	-0.01	*(.047)*
Hispanic	0.02	*(.031)*	-0.30 ***	*(.027)*	-0.07 **	*(.027)*	-0.22 ***	*(.064)*	-0.26 ***	*(.043)*
Other Race	-0.27 ***	*(.029)*	-0.41 ***	*(.028)*	-0.22 ***	*(.027)*	-0.55 ***	*(.062)*	-0.45 ***	*(.045)*
Socio-Economic Status	0.48 ***	*(.011)*	0.16 ***	*(.011)*	0.31 ***	*(.011)*	0.51 ***	*(.021)*	0.13 ***	*(.018)*
Testscore	0.81 ***	*(.012)*	0.40 ***	*(.011)*	0.5 ***	*(.011)*	0.83 ***	*(.021)*	0.32 ***	*(.017)*
Midwest	-0.01	*(.028)*	0.13 ***	*(.027)*	0.06 *	*(.027)*	0.11 *	*(.054)*	0.10 **	*(.047)*
South	-0.18 ***	*(.027)*	-0.03	*(.026)*	-0.15 ***	*(.025)*	-0.07	*(.051)*	0.17 ***	*(.044)*
West	-0.22 ***	*(.031)*	-0.12 ***	*(.029)*	-0.13 ***	*(.029)*	0.03	*(.067)*	0.18 ***	*(.051)*
Suburban	0.05 *	*(.023)*	0.13 ***	*(.021)*	0.09 ***	*(.021)*	0.09 **	*(.044)*	0.14 ***	*(.034)*
Rural	-0.01	*(.029)*	0.14 ***	*(.028)*	-0.01	*(.027)*	0.16 ***	*(.056)*	0.11 **	*(.047)*
Catholic School	0.78 ***	*(.038)*	0.22 ***	*(.037)*	0.49 ***	*(.036)*	0.95 ***	*(.065)*	0.4 ***	*(.065)*
Non-Catholic Private Sch	-0.09 **	*(.037)*	-0.20 ***	*(.039)*	-0.24 ***	*(.035)*	-0.09	*(.060)*	-0.57 ***	*(.073)*

Notes:

1) The parameter on log hourly wages (relating wage utility to non-pecuniary utility) is 1.37, with SE of *(.002)* .

2) The variance of the normal wage error terms is estimated to be 0.16, with a SE of *(.001)* .

4)*,**,*** denote 90%, 95%, and 99% statistical significance respectively.

5) Total # Observations is 16,200.

6) Standard errors (SE) are calculated using the covariance of the parameter estimate scores, following Train (2003).

Table C2: Additional PSE Structural Parameters

Variable	1-yr Trade School Estimate		*SE*	2-yr CC Estimate		*SE*	4-yr University Estimate		*SE*
1. PSE Instruments									
% Prev Students Enter Labor Market (0-5 Scale)	-0.01		*(.043)*	-0.04	*	*(.020)*	-0.18	***	*(.021)*
% Prev Students Attend 2yr College (0-5 Scale)	0.10	***	*(.038)*	0.18	***	*(.018)*	-0.16	***	*(.021)*
% Prev Students Attend 4yr College (0-5 Scale)	0.05		*(.032)*	-0.11	***	*(.015)*	0.51	***	*(.018)*
% Students Attend College Fairs	-0.04		*(.031)*	0.01		*(.014)*	0.03		*(.016)*
% Students in College App Prog (0-5 Scale)	0.01		*(.027)*	0.04	***	*(.012)*	0.18	***	*(.014)*
2. Personal Characteristics									
Male	-0.82	***	*(.106)*	0.08	**	*(.037)*	-1.09	***	*(.041)*
Black	0.54	***	*(.104)*	0.09	*	*(.054)*	2.12	***	*(.051)*
Hispanic	-0.03		*(.103)*	-0.04		*(.049)*	0.61	***	*(.050)*
Other Race	-0.01		*(.104)*	-0.09	*	*(.046)*	1.94	***	*(.052)*
Socio-Economic Status	0.17	***	*(.045)*	0.45	***	*(.019)*	2.15	***	*(.022)*
Testscore	0.13	**	*(.055)*	0.64	***	*(.021)*	3.31	***	*(.024)*
Midwest	0.11		*(.106)*	-0.09	*	*(.048)*	-0.90	***	*(.054)*
South	0.11		*(.101)*	-0.17	***	*(.046)*	-0.68	***	*(.051)*
West	0.09		*(.114)*	-0.08		*(.052)*	-0.92	***	*(.058)*
Suburban	-0.04		*(.077)*	0.10	***	*(.037)*	-0.74	***	*(.038)*
Rural	-0.07		*(.102)*	-0.01		*(.049)*	-0.91	***	*(.049)*
Catholic School	0.26	*	*(.132)*	0.75	***	*(.060)*	2.71	***	*(.080)*
Non-Catholic Private School	0.01		*(.134)*	0.04		*(.060)*	1.29	***	*(.082)*
Constant	-2.67	***	*(.304)*						
3. Previous Education									
Academic	-0.43	***	*(.119)*						
Business Vocational	-0.12		*(.315)*						
Trade Vocational	-1.31	***	*(.542)*						
Other Curriculum	0.20	**	*(.098)*						
GED	0.19		*(.157)*						

Notes:
2) *,**,*** denote 90%, 95%, and 99% statistical significance respectively.
3) Total # Observations is 16,200.
4) Standard errors (SE) are calculated using the covariance of the parameter estimate scores, following Train (2003).

Table C3: Additional HS Education Structural Parameters

Variable	Academic Estimate		Academic SE	General Ed Estimate		General Ed SE	Business Voc Estimate		Business Voc SE	Trade Voc Estimate		Trade Voc SE	Other Estimate		Other SE	GED Estimate		GED SE
1. HS Education Instruments																		
# Voc Teachers per 100 Students	-		-	-		-	-0.01		*(.144)*	-0.02		*(.144)*	-0.02		*(.094)*	-		-
% Stu Free / Reduced Price Lunch	0.12		*(.116)*	-0.28	***	*(.084)*	-0.01		*(.149)*	-0.02		*(.148)*	-0.23	**	*(.110)*	-0.44	***	*(.151)*
Admission Based on Geography	7.01	***	*(.110)*	6.86	***	*(.101)*	6.88	***	*(.129)*	6.76	***	*(.124)*	6.74	***	*(.110)*	8.09	***	*(.258)*
Stu Infl on Course Selection (0-3 Scale)	1.77	***	*(.053)*	1.8	***	*(.050)*	1.83	***	*(.060)*	1.80	***	*(.061)*	1.87	***	*(.053)*	1.81	***	*(.118)*
% Prev Stu Enter Labor Market (0-5 Scale)	2.75	***	*(.053)*	2.77	***	*(.050)*	2.90	***	*(.058)*	2.79	***	*(.057)*	2.77	***	*(.052)*	2.62	***	*(.119)*
% Prev Stu Attend 4yr Col (0-5 Scale)	-1.45	***	*(.055)*	-1.46	***	*(.053)*	-1.40	***	*(.061)*	-1.41	***	*(.059)*	-1.42	***	*(.055)*	-1.66	***	*(.124)*
% Prev Stu Attend 2yr Col (0-5 Scale)	4.35	***	*(.042)*	4.31	***	*(.040)*	4.33	***	*(.047)*	4.29	***	*(.046)*	4.30	***	*(.042)*	4.08	***	*(.092)*
GED Offered	-		-	-		-	-		-	-		-	-		-	2.15	***	*(.210)*
2. Personal Characteristics																		
Male	-0.81	***	*(.065)*	-0.48	***	*(.057)*	-0.53	***	*(.087)*	0.90	***	*(.092)*	-0.47	***	*(.068)*	1.27	***	*(.111)*
Black	0.40	***	*(.101)*	0.43	***	*(.085)*	0.27	**	*(.119)*	-0.31	**	*(.123)*	-0.04		*(.097)*	-0.76	***	*(.145)*
Hispanic	-0.07		*(.098)*	-0.12		*(.083)*	-0.67	***	*(.125)*	-0.80	***	*(.119)*	-0.47	***	*(.096)*	-0.59	***	*(.135)*
Other Race	0.46	***	*(.087)*	-0.12		*(.077)*	-0.23	*	*(.117)*	-0.53	***	*(.111)*	-0.51	***	*(.092)*	-0.55	***	*(.132)*
Socio-Economic Status	1.10	***	*(.038)*	0.91	***	*(.034)*	0.79	***	*(.048)*	0.71	***	*(.049)*	0.78	***	*(.039)*	0.69	***	*(.059)*
Testscore	2.34	***	*(.042)*	1.23	***	*(.035)*	1.15	***	*(.051)*	0.75	***	*(.048)*	0.75	***	*(.039)*	0.87	***	*(.064)*
Midwest	-0.11		*(.092)*	0.29	***	*(.080)*	0.23	*	*(.117)*	0.06		*(.111)*	0.49	***	*(.096)*	0.35	**	*(.138)*
South	-0.17	*	*(.082)*	-0.27	***	*(.072)*	0.04		*(.111)*	-0.67	***	*(.104)*	-0.18	*	*(.090)*	-0.03		*(.126)*
West	0.47	***	*(.097)*	0.89	***	*(.084)*	0.13		*(.143)*	-0.02		*(.129)*	0.85	***	*(.102)*	0.57	***	*(.152)*
Suburban	0.05		*(.071)*	0.35	***	*(.062)*	0.54	***	*(.092)*	0.61	***	*(.093)*	0.52	***	*(.073)*	0.59	***	*(.105)*
Rural	-0.42	***	*(.100)*	-0.05		*(.088)*	0.19		*(.118)*	0.08		*(.121)*	0.24	**	*(.099)*	0.15		*(.140)*
Catholic School	14.81	***	*(.224)*	15.29	***	*(.218)*	15.64	***	*(.263)*	13.82	***	*(.294)*	13.48	***	*(.251)*	14.74	***	*(.454)*
Non-Catholic Private School	3.67	***	*(.111)*	4.65	***	*(.089)*	4.19	***	*(.197)*	2.95	***	*(.263)*	3.12	***	*(.151)*	5.27	***	*(.157)*
Constant																24.0	***	*(.768)*

Notes:
1)*,**,*** denote 90%, 95%, and 99% statistical significance respectively.
2) Total # Observations is 16,200.
3) Standard errors (SE) are calculated using the covariance of the parameter estimate scores, following Train (2003).

D. Policy Simulations

D.1 Federal Vocational Offering Requirements

The structural estimates suggest that both business and trade high school vocational education are beneficial for certain non-college bound students. The first policy simulation investigates whether expanding vocational curricula offerings would incentivize more students to concentrate in these curricula and improve their later life outcomes. Specifically, it simulates the effects of a federal mandate requiring business and trade CTE to be taught on-site in every high school nationwide.[6]

The results of this simulation are shown in Column 3 of Table 6. The simulation suggests that this policy would increase the percent of students who take high school vocational curricula from 8.4% to 13.2% of the population. This change in high school curricula choice, in turn, would cause a few additional individuals to complete two-year community college degrees (11.7% → 12.0%) and a few less individuals to work in unskilled labor occupations at age 26 (4.2% → 4.1%). The simulation predicts that the policy would slightly increases the average log wages of individuals who switch their high school curricula to a vocational curricula (by 1.8%) and slightly increases their average lifetime utility (by 0.7%). Overall, the simulation leads to positive but relatively minor long-term effects on individuals' education and labor market outcomes.

D.2 Vocational Certificates in High School

The second simulation investigates the effects of incorporating vocational certifications into high school vocational curricula. Historically, vocational high school education in the United States has not included industry certification exams or certificate conferral – students have had to take relevant certification exams after graduating from high school, by attending one-year PSE trade schools or taking exams independently, to become certified (Castellano et al., 2005). Over the last few years, and after students in the ELS:2002 sample had graduated from high school, however, there has been a notable increase in the number of high school vocational programs that confer vocational certifications due to the re-authorization of the Carl D. Perkins Career and

[6] Note that there is a cost to schools of increasing CTE course offerings. First, many schools do not have teachers who are prepared to teach CTE courses. In these schools, there are additional costs to either training teachers to instruct these courses, or hiring new teachers who are prepared to teach them. In addition, cost analyses conclude that CTE courses are more expensive to offer than general education courses due to small CTE class sizes (Bishop and Mane, 2004; Rosa, 2009). While increasing CTE class sizes is one potential solution to this cost difference, it is unclear how increasing CTE class size may affect the quality and returns of the curricula.

Technical Education Act of 2006 (U.S. Department of Education, 2013). This policy simulation investigates the effects of this change on students' high school education, PSE attainment, and labor market outcomes.[7]

To conduct this simulation, I update the model so that an individual who completes a high school trade or business vocational curriculum immediately receives a one-year PSE trade school degree. Additionally, the individual receives the non-pecuniary utility associated with attending a one-year PSE trade school during her fourth year of high school in addition to the non-pecuniary utility she receives from her high school field choice that year.[8]

The results of this policy simulation are presented in Column 4 of Table 6. This policy incentivizes additional students to concentrate in a trade vocational curriculum (from 4.3% to 7.2% of the population) as it allows them to receive both a high school diploma and an industry certification concurrently. However, fewer individuals receive terminal degrees from a community college (11.7% → 9.1%) or a four-year university (48.2% → 47.1%), because fewer individuals take academic and general education courses in high school. Additionally, the policy leads to more individuals working in skilled non-manual labor occupations (21.3% → 21.7%) and skilled manual labor occupations (17.7% → 18.2%) and less individuals working in unskilled occupations (4.2% → 3.9%) or choosing not to work (13.9% → 13.6%) at age 26. Among individuals who change their behavior, average log wages increase substantially (by 9.1%) as does expected lifetime utility (by 2.2%). Overall, the simulation predicts that incorporating vocational certifications into high school vocational curricula would have substantial positive effects on student labor market outcomes.

D.3 Free Community College

An additional simulation leverages the model to explore the effects of an indirectly-related policy: providing free community college for all United States high school graduates. A version of this policy was proposed by President Barack Obama in 2015 and, more recently, several

[7] Note that an implementation evaluation of this change conducted by RTI International (Klein et al., 2014) found that the changes were broadly implemented but that the characteristics and quality of the updated programs varied considerably. States and local school districts reported that this variation was due to challenges with the implementation related to a lack of funds, staff, time, and administrative guidance.

[8] Note that this specification assumes that the returns to high school CTE degrees and vocational certifications are driven by the knowledge a student learns and the degrees that are conferred as opposed to the signaling value of pursuing each degree separately. If the latter is true the results of this policy simulation would be upward biased.

presidential candidates including Senator Bernie Sanders, Senator Hillary Clinton, and President Joe Biden (Obama, 2015; Sanders, 2016; Clinton, 2016; Biden, 2020).[9]

To conduct this simulation, I decrease the cost of attending community college in the model by the average cost of a year of community college in 2004 ($2,700). Practically, I convert this yearly cost into a log hourly cost for someone who works a normal 40-hour workweek. Then, I multiply this value by my estimate for φ (1.37), the parameter that relates pecuniary wage utility to non-pecuniary utility in the model. Finally, I add this value to the total utility an individual receives from attending community college each year. While the monetary cost of community college is the same for all individuals, the non-pecuniary utility associated with this cost is likely higher for poorer students than for richer students due to a diminishing marginal utility of wealth. To incorporate this difference, I vary the reduction in the non-pecuniary cost of community college across individuals based on their reported socio-economic status. For the simulation presented here, the individual with the highest socio-economic status in the sample receives no additional non-pecuniary utility from attending community college while the individual with the lowest socio-economic status in the sample receives twice the average non-pecuniary cost-savings when attending community college.[10]

The results of this simulation are presented in Table D1. Decreasing the cost of community college causes substantially more students to graduate from community college (from 11.7% to 27.0% of the population). The simulation predicts that more students would concentrate in general education courses in high school (44.9% → 46.0%) – as completing a high school general education curriculum improves the utility of attending community college – largely at the expense of taking academic courses (22.8% → 21.2%). In addition, the simulation predicts that fewer individuals would drop out of high school and have no high school degree by age 26 (4.9% → 4.2%) as high school graduation is required to attend community college.

The substantial increase in community college graduation is driven by two subsets of the population: individuals who previously did not attend any PSE institutions, and individuals who

[9] Bernie Sanders and Hillary Clinton proposed plans that, in addition to providing free tuition to community colleges, also provide free tuition to certain four-year colleges and universities and include additional debt relief. This policy simulation does not include these additions and focuses on the effects of the central plan to provide free tuition to community colleges for all U.S. high school graduates.

[10] In reality a subset of low socio-economic status individuals receive Pell Grants that decrease the cost of community college to close to zero. A question of future work involves incorporating these Pell Grants into the simulation by holding the cost of community college fixed for the subset of students in the population who are eligible to receive these grants.

previously attended four-year universities but now instead attend community colleges due to their decreased cost. Overall, the simulation predicts a decrease in the number of individuals who graduate from four-year universities by age 26 (48.2% → 44.8%). However, recall that the model does not account for transferability of community college credits to four-year universities. As such, this prediction is likely a lower bound of the effect of this policy on four-year university graduation. Using a back-of-the-envelope calculation, and assuming the community college to four-year university transfer rate remains unchanged at 20% (Hossler et al., 2012), the simulation instead predicts a small net increase in the percent of individuals who graduate from (or are currently attending) four-year universities at age 26 (59.5% → 60.8%).

Overall, the simulation predicts a decrease in average wages at age 26 (by -3.7%) and an increase in average welfare (by 0.8%) among individuals who change their behavior or attend community college, with the increase in welfare driven by the decreased financial burden of community college. Under a back-of-the-envelope calculation that assumes a 20% community college to four-year university transfer rate these numbers increase considerably: under this assumption the simulation predicts an increase in average wages (by 0.5%) and an increase in average welfare (by 1.0%) among individuals who change their behavior or attend community college. These results suggest that the positive labor market effects of this policy will be driven by the number of new community college graduates who transfer to, and graduate from, four-year universities.

Finally, note that these welfare estimates do not account for the costs of the policy. Under an assumption that these costs are bourn equally by every individual in the population this policy would decrease average welfare. Under an assumption that the costs are disproportionately bourn by individuals with the highest socio-economic statuses in the population, who have the lowest marginal utilities of wealth, this policy could increase average welfare in the population even after accounting for cost. The extent of this increase (or decrease) in welfare, however, depends on one's assumptions about relative social marginal welfare across individuals of different socio-economic status.

Table D1: Free CC Policy Simulation

Variable	Simulation No Policies	Free CC
i. HS Graduation Curriculum		
Academic	22.8%	-1.6%
Gen Ed	44.9%	1.1%
Bus Voc	4.1%	0.5%
Trade Voc	4.3%	0.2%
Other	15.4%	0.2%
GED	3.5%	0.3%
Never Graduate	4.9%	-0.7%
ii. PSE Degrees		
HS Grad Or Less	38.2%	-5.0%
1-yr Trade	6.8%	-1.9%
2-yr Community College	11.7%	15.3%
4-yr University	48.2%	-3.4%
iii. Employment Age 26		
Professional	28.5%	-0.4%
Skilled Manual Labor	17.7%	-0.5%
Skilled Non-Manual Labor	21.3%	-0.6%
Skilled Other	3.0%	-0.1%
Unskilled	4.2%	-0.2%
Unemployed	13.9%	-0.2%
Attending PSE	11.3%	2.0%
iv. Wages Age 26		
Average (ln) Hourly Wage	2.42	-3.7%
v. Utility		
Lifetime Utility	150.0	0.8%
Realized Utility Ages 16-26	104.7	0.9%
Expected Utility Ages 27+	45.3	0.4%
vi. % Changed Observations		
% Changed Average (ln) Hourly Wage		13%
% Changed Realized Utility Ages 16-26		65%
% Changed Expected Utility Ages 27+		56%

Notes:
Column (1) displays displays simulated outcomes, given the model, structural parameters, and initial conditions in the ELS:2002 data set. Column (2) panels i-iii display the percentage point difference between the baseline simulation in column (1) and the simulated outcomes for the "free community college" policy simulation. Column (2) panels iv-v display the percent difference between the baseline simulation in column (1) and the average simulated outcomes for the "free community college" policy simulation, among individuals whose outcome value changed between the baseline outcome and simulated outcome, and in the case of wages conditional on an occupation being chosen in 2012 in both simulations. Panel vi, "% Changed Observations", denotes the percent of individuals who meet these conditions. PSE degrees are cummulative: An individal in the sample can have multiple types of PSE credentials. Hence the total number of PSE degrees in each column can be higher than the number of individuals in the sample. Total # Observations is 16,200.

E. Estimation

E.1 Expected Lifetime Utility

Define $\tilde{\varepsilon}_{it}$ and ε_{it} (without superscripts) as the vectors of wage time-specific error terms and non-pecuniary time-specific error terms, respectively, for all choices for individual *i* in period *t*. Define S_{it} as the state vector for individual *i* at the start of period *t*, which consists of relevant time-invariant characteristics about the individual $(X_i, \tilde{u}_i^k, u_i^k)$, vectors of past education and employment decisions $(F_{it}, H_{it}, N_{it}, P_{it}, O_{it})$, and vectors of current period time-specific stochastic error terms $\tilde{\varepsilon}_{it}$ (wage utility) and ε_{it} (non-pecuniary utility). Time-invariant characteristics about the individual $(X_i, \tilde{u}_i^k, u_i^k)$ do not change in S_{it} over time. The variables F_{it+1}^k and N_{it+1}^k increase by one with certainty every year the individual chooses to attend high school in a specific field and chooses to attend a specific type of PSE institution. H_{it+1}, P_{it+1}, and O_{it+1} change, as defined in Section 3.1, when the individual graduates from high school, graduates from each type of PSE institution, and works in a particular field and gains occupation-specific human capital.[11]

Denote individual *i*'s choice in period *t* as k_{it}. I define the transition of the state vector described in the preceding paragraph as

$$S_{it+1} = G(S_{it}, k_{it}, \psi_{it}, \tilde{\varepsilon}_{it+1}, \varepsilon_{it+1}) \quad . \tag{3}$$

Note that today's choice between available education and labor market options (k_{it}) affects future choices $(k_{i\tau}, \tau > t)$ by increasing the stock values of $F_{it}, H_{it}, N_{it}, P_{it}$, and O_{it} for every future period $\tau = t+1, t+2, \dots, T$. These increased stock values affect the value of utility for each choice in every future period $\tau = t+1, t+2, \dots, T$.

An individual chooses between her education and employment options in each period *t* to maximize her expected lifetime utility between the current period and retirement at age 65 $(t = T)$. The individual's expected lifetime utility (i.e., value function) at the start of period *t* can be written as

$$V_{it}(S_{it}) = \max_{\{k\}} \left[U_{it}^k(S_{it}) + E\left(\sum_{\tau=t+1}^{T} \delta^{\tau-t} \max_{\{\kappa\}} \; U_{i\tau}^{\kappa}(S_{i\tau}) \right) \right]$$

where δ is the discount factor, $U_{it}^k(S_{it})$ is the current period utility from choosing option *k* given

[11] There are 3,360 possible states of occupation-specific human capital and educational attainment in the model, comprised of 15 states of occupation-specific human capital and 224 states of education experience. The states of education experience are comprised of 56 states of high school education experience prior to high school graduation, 144 states of HS degree and PSE experience after high school graduation but prior to four-year university graduation, and 24 different states of HS degree and PSE degree attainment after four-year university graduation.

state vector S_{it}, and S_{it} follows the transition of the state vector noted in Equation 3. The mean $E(\cdot)$ is over the joint distribution of future error terms $\psi_{i\tau}$, $\tilde{\varepsilon}_{i\tau}$, $\varepsilon_{i\tau}$ for every period $\tau = t+1, t+2, \dots, T$.

Define $\bar{S}_{it}$ as the pre-period state, prior to the start of period *t*, which consists of everything in state vector S_{it} except period *t* error term vectors $\tilde{\varepsilon}_{it}$ and ε_{it}. The expected value of lifetime utility from period *t* until retirement, prior to realizing the error term vectors $\tilde{\varepsilon}_{it}$ and ε_{it} that are drawn at the start of period *t*, can be written as

$$V_{it}^{*}(\bar{S}_{it}) = E\left[\sum_{\tau=t}^{T} \delta^{\tau-t} \max_{\{k\}} U_{i\tau}^{k}(S_{i\tau})\right]$$

where the mean $E(\cdot)$ is over the joint distribution of future error terms $\psi_{i\tau}$, $\tilde{\varepsilon}_{i\tau}$, $\varepsilon_{i\tau}$ in every period $\tau = t, t+1, t+2, \dots, T$. Thus, the net present value of choosing choice k today, after realizing today's error term vectors $\tilde{\varepsilon}_{it}$ and ε_{it}, can be rewritten using Bellman's equation as

$$V_{it}^{k}(S_{it}) = U_{it}^{k}(S_{it}) + \delta V_{it+1}^{*}(\bar{S}_{it+1}) \quad .$$

Note that tomorrow's pre-period state ($\bar{S}_{it+1}$) is determined based on today's state vector (S_{it}) and today's choice (k_{it}) as noted in Equation 3. Because the non-pecuniary error terms for each choice (ε_{it}^{k}) are distributed $iid\ EV(0,1)$, the expected value of lifetime utility from period *t* until retirement, prior to realizing today's time-specific error terms, has the following closed-form solution:

$$V_{it}^{*}(\bar{S}_{it}) = \int ln\left(\textstyle\sum_j \exp\{\bar{V}_{it}^{j}(S_{it})\}\right) f(\tilde{\varepsilon}_{it}) d\tilde{\varepsilon}_{it} \qquad (4)$$

$$\text{where } \bar{V}_{it}^{j}(S_{it}) = V_{it}^{j}(S_{it}) - \varepsilon_{it}^{j} \quad .$$

The integral over $\tilde{\varepsilon}_{it}$ corresponds to integrating over each of the normal $\tilde{\varepsilon}_{it}^{k}$ error terms associated with wages in each of the five occupations.[12]

E.2 Likelihood Function

I estimate the parameters in the model using maximum simulated likelihood estimation. The likelihood function is constructed as described below. First, define an individual's realized

[12] The construction of the value function is similar to the derivation used in other dynamic discrete choice models such as Keane and Wolpin (1997) and Chan (2013).

log-wage offer in occupation *k* in period *t* as $\widehat{w}_{it}^{k}$, define d_{wit}^{k} as a binary variable equal to one if $\widehat{w}_{it}^{k}$ is observed in the data set, and define $\omega_{it} = (\widehat{w}_{it}^{1}, d_{wit}^{1}, \widehat{w}_{it}^{2}, d_{wit}^{2}, \dots, \widehat{w}_{it}^{5}, d_{wit}^{5})$. Note that each ω_{it} contains at most one non-zero d_{wit}^{k} as I observe at most one log-wage offer in the data set for an individual each period.

Recall that each pre-period state $\bar{S}_{it}$ includes the personal characteristics of the individual (X_i), the unobserved heterogeneity type of the individual (u_i), the previous high school and post-secondary education experience of the individual ($F_{it}, H_{it}, N_{it}, P_{it}$), and the previous human capital accumulation of the individual (O_{it}). Also, recall that each state vector S_{it} includes $\bar{S}_{it}$ as well as current period utility and log-wage error terms $\tilde{\varepsilon}_{it}$ and ε_{it}. Define the expected value of log wages in occupation *k* in period *t* as

$$E\left[w_{it}^{k}(\bar{S}_{it})\right] = w_{it}^{k}(S_{it}) - \tilde{\varepsilon}_{it}^{k}$$

and define an individual's residual log-wage error term associated with realized log-wage offer $\widehat{w}_{it}^{k}$ as

$$\hat{\tilde{\varepsilon}}_{it}^{k} = \widehat{w}_{it}^{k} - E\left[w_{it}^{k}(\bar{S}_{it})\right] \quad . \tag{5}$$

Finally, define $f\left(\tilde{\varepsilon}_{it} \backslash \tilde{\varepsilon}_{it}^{k} \mid \hat{\tilde{\varepsilon}}_{it}^{k}\right)$ as the joint density function of the log-wage error terms for every occupation except occupation *k*, conditional on an observed residual log-wage error term for occupation *k*. As each $\tilde{\varepsilon}_{it}^{k}$ is assumed to be *iid*, the joint density of the other $\tilde{\varepsilon}_{it}^{k}$'s does not depend on the value of the residual $\hat{\tilde{\varepsilon}}_{it}^{k}$. That is, $\tilde{\varepsilon}_{it} \backslash \tilde{\varepsilon}_{it}^{k} \mid \hat{\tilde{\varepsilon}}_{it}^{k} \sim N(0, \sigma_{\tilde{\varepsilon}}^{2} I)$, where I is a four-by-four identity matrix corresponding to the four other occupations in period *t*.

Recall that $\bar{V}_{it}^{k}(S_{it})$ is a function of $w_{it}^{k}(S_{it})$ which is itself a function of $\tilde{\varepsilon}_{it}^{k}$. Because the non-pecuniary error terms for each choice (ε_{it}^{k}) are distributed $iid\ EV(0,1)$, the conditional likelihood that individual *i*, with pre-period state $\bar{S}_{it}$, chose choice *k* in period *t* is

$$L_{cit}^{k}(\bar{S}_{it}, \omega_{it}) = \int \frac{\exp\{\bar{V}_{it}^{k}(S_{it})\}}{\sum_j \exp\{\bar{V}_{it}^{j}(S_{it})\}} f\left(\tilde{\varepsilon}_{it} \backslash \tilde{\varepsilon}_{it}^{k} \mid \hat{\tilde{\varepsilon}}_{it}^{k}\right) d\tilde{\varepsilon}_{it} \backslash \tilde{\varepsilon}_{it}^{k} \quad if \quad d_{wit}^{k} = 1\,, \tag{6}$$

$$L_{cit}^{k}(\bar{S}_{it}, \omega_{it}) = \int \frac{\exp\{\bar{V}_{it}^{k}(S_{it})\}}{\sum_j \exp\{\bar{V}_{it}^{j}(S_{it})\}} f(\tilde{\varepsilon}_{it}) d\tilde{\varepsilon}_{it} \quad if \quad d_{wit}^{k} = 0\,,$$

where $\quad \tilde{\varepsilon}_{it}^{k} = \hat{\tilde{\varepsilon}}_{it}^{k} \quad \text{iff} \quad d_{wit}^{k} = 1 \quad .$

Note that ω_{it} has two effects on the likelihood function when a wage is observed ($d_{wit}^{k} = 1$). First, the corresponding residual log-wage error term ($\hat{\tilde{\varepsilon}}_{it}^{k}$) is directly inserted into the likelihood function.

Second, $\hat{\tilde{\varepsilon}}_{it}^{k}$ affects the conditional joint distribution of the remaining unobserved error terms ($f\left(\tilde{\varepsilon}_{it} \backslash \tilde{\varepsilon}_{it}^{k} \mid \hat{\tilde{\varepsilon}}_{it}^{k}\right)$), which is integrated over to calculate the likelihood function.

Every period that a log wage is observed a wage likelihood can be calculated. Because each log-wage error term is distributed $iid\ N(0, \sigma_{\tilde{\varepsilon}}^{2})$, the conditional likelihood that a particular log wage was offered in occupation *k* in period *t,* given pre-period state $\bar{S}_{it}$, is

$$L_{wit}^{k}(\bar{S}_{it}, \omega_{it}) = \left(\frac{1}{\sigma_{\tilde{\varepsilon}}}\right) \phi\left(\frac{\hat{\tilde{\varepsilon}}_{it}^{k}}{\sigma_{\tilde{\varepsilon}}}\right) \qquad \text{iff} \quad d_{wit}^{k} = 1 \ .$$

Thus, the total conditional likelihood contribution for individual *i* in period *t*, given a particular pre-period state $\bar{S}_{it}$ and observed wage vector ω_{it}, is

$$L_{it}^{k}(\bar{S}_{it}, \omega_{it}) = L_{cit}^{k}(\bar{S}_{it}, \omega_{it}) L_{wit}^{k}(\bar{S}_{it}, \omega_{it}) \quad \text{if} \ \ d_{wit}^{k} = 1 \ , \tag{7}$$

$$L_{it}^{k}(\bar{S}_{it}, \omega_{it}) = L_{cit}^{k}(\bar{S}_{it}, \omega_{it}) \qquad \text{if} \ \ d_{wit}^{k} = 0 \ .$$

Define the path of choices over the individual's lifetime as $K_{pi} = \{k_{i1}, k_{i2}, \ldots, k_{iT}\}$, the associated pre-period state path over the individual's lifetime as $\bar{S}_{pi} = \{\bar{S}_{i1}, \bar{S}_{i2}, \ldots, \bar{S}_{iT}\}$, and the path of observed wages over the individual's lifetime as $\omega_{pi} = \{\,\omega_{i1}, \omega_{i2}, \ldots, \omega_{iT}\,\}$.[13] The conditional lifetime likelihood function for individual *i* is a function of the path of choices over her lifetime (K_{pi}), the associated pre-period states over her lifetime ($\bar{S}_{pi}$), and the observed wage information over her lifetime (ω_{pi}):

$$L_{li}\left(K_{pi}, \bar{S}_{pi}, \omega_{pi}\right) = \prod_{t=1}^{T} L_{it}^{k}(\bar{S}_{it}, \omega_{it}) \quad .$$

E.3 Unobserved Events

However, I do not always observe K_{pi} and $\bar{S}_{pi}$ because I do not observe the choices an individual makes during periods where information is missing in the data set (when k_{it} is unknown)

[13] Note that a choice path (K_{pi}) can be mapped to multiple state vector paths ($\bar{S}_{pi}$), and that a state vector path ($\bar{S}_{pi}$) can be mapped to multiple choice paths (K_{pi}). For example, choosing to work in a professional occupation in period *t* (k_{it} = *Professional*) can have two possible effects on $\bar{S}_{it+1}$ depending on whether or not occupation-specific human capital (O_{it}) is gained. Conversely, the state space transition of $\bar{S}_{it} = \bar{S}_{it+1}$ can be caused by both choosing not to be employed (k_{it} = *Not Employed*) or choosing to work in the professional field and not gaining occupation-specific human capital (k_{it} = *Professional*, $\psi_{it} = 0$).

and do not observe when an individual gains occupation-specific human capital. Define the path of occupation-specific human capital over an individual's lifetime as $O_{pi} = \{O_{i1}, O_{i2}, \dots, O_{iT}\}$, and note that $O_{pi} \in \bar{S}_{pi}$. Define d_{it}^{o} as a binary variable equal to one if the individual's choice in period t (k_{it}) is observed in the data set, T_i^o as the set of all time periods for which d_{it}^{o} = 1 for individual i, and K_{pi}^{o} as the set of all k_{it}'s for which d_{it}^{o} = 1 for individual i. Note that, for every possible choice path (K_{pi}) and every possible occupation-specific human capital accumulation path (O_{pi}), I can calculate the individual's associated lifetime likelihood ($L_{li}\left(K_{pi}, \bar{S}_{pi}, \omega_{pi}\right)$). The conditional lifetime likelihood contribution of an individual with missing information can then be calculated as a weighted sum of the conditional lifetime likelihood functions for each possible path of education and employment that could have taken place for the individual:

$$L_{ui}\left(u_i, X_i, K_{pi}^{o}, \omega_{pi}\right) = \sum_{\bar{S}_{pi}} P\left(\bar{S}_{pi} \middle| K_{pi}^{o}\right) \prod_{T_i^o} L_{it}^{k}(\bar{S}_{it}, \omega_{it})$$

where the summation is over all possible state paths $\bar{S}_{pi}$ such that u_i,$X_i \in \bar{S}_{pi}$, and $P\left(\bar{S}_{pi} \middle| K_{pi}^{o}\right)$ is the probability that pre-period state path $\bar{S}_{pi}$ occurred given observable choices K_{pi}^{o}.

Next, note that the probability the individual chose choice k in period t when k is unobserved ($d_{it}^{o} = 0$) is also $L_{it}^{k}(\bar{S}_{it}, \omega_{it})$ and that the probability the individual accumulated human capital in period t if she worked and had education level e is θ_e. As such, the conditional lifetime likelihood contribution for individual i with unobserved heterogeneity type u_i and personal characteristics X_i can be rewritten as

$$L_{ui}\left(u_i, X_i, K_{pi}^{o}, \omega_{pi}\right) = \sum_{\bar{S}_{pi}} \sum_{K_{pi}} P\left(\bar{S}_{pi} \middle| K_{pi}\right) L_{li}\left(K_{pi}, \bar{S}_{pi}, \omega_{pi}\right)$$

where the second summation is over all K_{pi} such that $K_{pi}^{o} \in K_{pi}$. Note that $P\left(\bar{S}_{pi} \middle| K_{pi}\right)$ is comprised entirely of a product of θ_e's and $[1 - \theta_e]$'s based on whether O_{it}^{k} (occupation-specific human capital) increased each period the individual worked along choice path K_{pi}.[14] Also, note that $L_{li}\left(K_{pi}, \bar{S}_{pi}, \omega_{pi}\right)$ is a product of the conditional period likelihood contributions ($L_{it}^{k}(\bar{S}_{it}, \omega_{it})$) for individual i for every period $t = 1,2, \dots, T$. This includes the likelihood contributions for

[14] For example, if an individual never graduated from high school and worked in a skilled manual labor job in every period t = 1,2,…,T, the probability that pre-period state path $\bar{S}_{pi}$ occurred in which no occupation-specific human capital was accumulated is $P\left(\bar{S}_{pi} \middle| K_{pi}\right)$ = $[1 - \theta_{noHS}]^{T}$.

periods where choice $k_{it} \in K_{pi}$ is observed (d_{kit} = 1) as well as the likelihood contributions for periods when choice $k_{it} \in K_{pi}$ is unobserved (d_{kit} = 0).

Finally, note that $L_{ui}(u_i, X_i, K_{pi}^o, \omega_{pi})$ is the lifetime likelihood contribution for an individual with unobserved heterogeneity type u_i. Since I do not observe whether the person is a type-one or type-two individual, the individual's overall lifetime likelihood function is the weighted sum of her type-one and type-two conditional lifetime likelihood functions, where the weights are the percentages of each type of individual in the population:

$$L_i(X_i, K_{pi}^o, \omega_{pi}) = \zeta L_{ui}(u_1, X_i, K_{pi}^o, \omega_{pi}) + (1-\zeta) L_{ui}(u_2, X_i, K_{pi}^o, \omega_{pi}) \quad .$$

The sample likelihood function (L) is the product of each sample member's individual likelihood contribution:

$$L = \prod_i L_i(X_i, K_{pi}^o, \omega_{pi}) \quad .$$

I estimate the model by selecting parameters that maximize this sample likelihood function.[15]

E.4 Simulation

Integrating over the distribution of each unknown wage error term $\tilde{\varepsilon}_{it}^k$ to calculate each $V_{it}^*(\bar{S}_{it})$ and $L_{cit}^k(\bar{S}_{it}, \omega_{it})$ function, as described in Equations 4 and 6, is computationally burdensome. Calculating the lifetime likelihood function for individual *i* for every possible choice path K_{pi} such that $K_{pi}^o \in K_{pi}$ and every pre-period state path $\bar{S}_{pi}$ such that $u_i, X_i \in \bar{S}_{pi}$ is also computationally burdensome. To simplify these calculations, simulation methods are used. First, 10 independent values for each wage error term ($\tilde{\varepsilon}_{it}^k$) are simulated using antithetic acceleration.[16] Define each simulated value of $\tilde{\varepsilon}_{it}^k$ as $\epsilon_{\xi it}^{vk}$, where the ξ subscript refers to the simulation number (ξ = 1,2,…,10) and the v superscript denotes that the value is used when simulating the value function ($V_{it}^*(\bar{S}_{it})$). Define a set of simulated values across all occupations k in period t as $\epsilon_{\xi it}^{v}$. The value of the integral in Equation 4 is approximated as

[15] Parameter values are chosen following the Berndt, Hall, Hall, and Hausman (1974) (BHHH) optimization algorithm. The covariance matrix of maximum simulated likelihood estimates is standard.

[16] Borsch-Supan and Hajivassiliou (1993) showed that 20 simulations without antithetic acceleration is a large enough number of simulations to produce consistent estimates. Geweke (1988) showed that antithetic acceleration reduces the sample size required to produce consistent estimates for an initial sample of 20 by at least a factor of four. As such, 10 simulations is large enough to construct consistent estimates of V_{it}^* and L_{ui}.

$$V^*_{it}(\bar{S}_{it}) \;\approx\; V^*_{\xi it}(\bar{S}_{it}, \epsilon^{v}_{\xi it}) \;=\; \left(\frac{1}{10}\right)\sum_{\xi=1}^{10}\left[ln\left(\sum_j \exp\{\bar{V}^j_{it}(S_{it})\}\right) \mid \tilde{\varepsilon}_{it} = \epsilon^{v}_{\xi it}\right] \quad .$$

Separately, 10 independent values of each ε^k_{it} and $\tilde{\varepsilon}^k_{it}$ are simulated for each available choice each period using antithetic acceleration, as are 10 independent values of ψ_{it} (related to human capital accumulation) each period. Define these simulated values as $\epsilon^{\varepsilon k}_{\xi it}$, $\epsilon^{\tilde{\varepsilon} k}_{\xi it}$, and $\epsilon^{\psi}_{\xi it}$, respectively, and collectively define a set of these simulated values across all occupations k in period t as $\epsilon_{\xi it}$. First, the value of L^k_{cit} in Equation 6, given pre-period state $\bar{S}_{it}$, is simulated as

$$L^k_{\xi cit}(\bar{S}_{it}, \omega_{it}, \epsilon_{\xi it}) = \frac{\exp\{\bar{V}^k_{it}(S_{it})\}}{\sum_j \exp\{\bar{V}^j_{it}(S_{it})\}} \quad ,$$

$$\text{where} \quad \tilde{\varepsilon}^k_{it} = \hat{\tilde{\varepsilon}}^k_{it} \quad \text{if} \quad d^k_{wit} = 1 \quad ,$$

$$\tilde{\varepsilon}^k_{it} = \epsilon^{\tilde{\varepsilon} k}_{\xi it} \quad \text{if} \quad d^k_{wit} = 0 \quad .$$

Following Equation 7, the simulated value of L^k_{it} is constructed as

$$L^k_{\xi it}(\bar{S}_{it}, \omega_{it}, \epsilon_{\xi it}) = L^k_{\xi cit}(\bar{S}_{it}, \omega_{it}, \epsilon_{\xi i}\ )L^k_{wit}(\bar{S}_{it}, \omega_{it}) \quad if \quad d^k_{wit} = 1 \,,$$

$$L^k_{\xi it}(\bar{S}_{it}, \omega_{it}, \epsilon_{\xi it}) = L^k_{\xi cit}(\bar{S}_{it}, \omega_{it}, \epsilon_{\xi i}\ ) \quad if \quad d^k_{wit} = 0 \,.$$

Next, when $d^o_{it} = 0$ (the choice in period t) is unobserved the predicted value of k, given pre-period state $\bar{S}_{it}$, is simulated as

$$k_{\xi it}(\bar{S}_{it}, \epsilon_{\xi it}) = \text{argmax}_k\{V^k_{it}(S_{it}) \mid \varepsilon_{it} = \epsilon^{\varepsilon}_{\xi it}\,, \tilde{\varepsilon}_{it} = \epsilon^{\tilde{\varepsilon}}_{\xi it}\} \quad .$$

Finally, human capital accumulation (O_{it}), given pre-period state $\bar{S}_{it}$, is simulated each period as

$$O^k_{\xi it+1}(\bar{S}_{it}, \epsilon_{\xi it}) = O^k_{it} + \epsilon^{\psi}_{\xi it} \quad \text{iff} \quad d^k_{it} = 1 \quad \& \quad \underset{j}{sum}(O^j_{it}) \le 2 \,,$$

$$O^k_{\xi it+1} = O^k_{it} \qquad \text{otherwise.}$$

Define $K_{\xi pi}$ as the simulated choice path that includes K_{opi} and a simulated $k_{\xi it}(\bar{S}_{\xi it}, \epsilon_{\xi it})$ in each period that choice k_{it} is unobserved, such that $\bar{S}_{\xi it} \in \bar{S}_{\xi pi}$, where $\bar{S}_{\xi pi} = \{\bar{S}_{\xi i}\ , \bar{S}_{\xi i2}, \ldots, \bar{S}_{\xi iT}\}$ is the associated simulated pre-period state path and each $\bar{S}_{\xi it}$ is constructed iteratively, starting from period one, based on $\bar{S}_{\xi it-1}$, $k_{it-1} \in K_{\xi pi}$, and $\epsilon^{\psi}_{\xi it-1}$ as defined in Equation 3. The conditional lifetime likelihood for a particular simulated choice path $K_{\xi pi}$, along pre-period state path $\bar{S}_{\xi pi}$, is

$$L_{\xi ui}\left(u_i, X_i, K_{opi}, \omega_{pi}, K_{\xi pi}, \bar{S}_{\xi pi}\right) = \prod_{T_i^o} L_{\xi it}^{k_{it}}\left(\bar{S}_{\xi it}, \omega_{it}, \epsilon_{\xi it}\right) \quad .$$

Recall that T_i^o is the set of all time periods for which the individual's choice was observed in the data set (i.e., all periods for which d_{it}^o = 1). The conditional lifetime likelihood function for individual *i* can be approximated as the average of 10 simulated conditional lifetime likelihoods:

$$L_{ui}\left(u_i, X_i, K_{pi}^o, \omega_{pi}\right) \approx \left(\frac{1}{10}\right) \sum_{\xi=1}^{10} L_{\xi ui}\left(u_i, X_i, K_{opi}, \omega_{pi}, K_{\xi pi}, \bar{S}_{\xi pi}\right) \quad .$$

E.5 Identification

Variation across individuals over time allows me to identify each of the parameters in the model. First, each parameter in the wage equation ($\tilde{\beta}_X^k, \tilde{\beta}_P^k, \tilde{\beta}_H^k, \tilde{\beta}_{PH}^k, \tilde{\beta}_O^k$) is identified by variation in choices and wages over time across individuals. For example, the effect of gender on wages in occupation *k* ($\tilde{\beta}_{X_{MALE}}^k$) is identified by co-variation between gender and wages in occupation *k* among individuals with otherwise equivalent pre-period states. The effect of occupation-specific human capital in occupation *j* on wages in occupation *k* ($\tilde{\beta}_{O_j}^k$) is identified by co-variation in the number of years worked in occupation *j* and wages in occupation *k* among individuals with otherwise equivalent pre-period states.

Each parameter in the non-pecuniary utility equation (β_X^k, β_H^k, φ) is also identified. For example, the utility effects of a business vocational high school curriculum on attending two-year community college ($\beta_{H_{BUS}}^{CC}$) is identified by co-variation in two-year community college attendance between individuals who completed a business vocational high school curriculum and individuals who completed a general education high school curriculum, among individuals that attended high schools with different vocational and PSE opportunities but with otherwise equivalent pre-period states. The total amount of additional utility (both pecuniary and non-pecuniary) males receive in occupation *k* ($\tilde{\beta}_{X_{MALE}}^k + \beta_{X_{MALE}}^k$) is identified by co-variation between gender and occupation choice among individuals with otherwise equivalent pre-period states. As the pecuniary portion of this utility ($\tilde{\beta}_{X_{MALE}}^k$) is identified from observed wages, the non-pecuniary portion of this utility ($\beta_{X_{MALE}}^k$) is identified by the difference between the estimate for total utility ($\tilde{\beta}_{X_{MALE}}^k + \beta_{X_{MALE}}^k$)

and pecuniary utility ($\tilde{\beta}^k_{X_{MALE}}$).[17]

Next, the distribution of unobserved heterogeneity values in the population ($\tilde{u}^k_2, u^k_2, \zeta$) is identified by variation across and persistence in individual choice paths and wages. For example, the magnitude of wage-related unobserved heterogeneity in the population in occupation k for type-two individuals ($\tilde{u}^k_2$) is identified by, for individuals across the sample with persistently higher or lower observed wages than average in occupation k over time, the extent to which their wages are higher and lower than average among individuals with otherwise equivalent pre-period states. The distribution of non-pecuniary-utility-related unobserved heterogeneity in the population in occupation k (u^k_2) is identified by, for individuals across the sample who persistently choose occupation k more than average, the extent of that persistence, among individuals with otherwise equivalent pre-period states and observed wages.

The variance of the normal wage error terms (σ^2_{ε}) is identified by the variation in residual log-wage error terms throughout the sample. The probabilities that individuals with different educational attainment levels accrue occupation-specific human capital from working ($\theta_{noHS}, \theta_{HS}, \theta_{1yr}, \theta_{CC}, \theta_{4yr}$) are identified by the rates at which observed wages in each occupation discretely increase from period to period for individuals with each level of educational attainment.

E.6 Structural vs. Non-Structural Estimates

The structural model has several advantages over non-structural models. First, by estimating a structural model, I can separately identify the intertemporal effects of education and labor market choices – how each choice impacts present and future utility – and the mechanisms underlying those effects. For example, by estimating a structural model, I can identify whether a student takes high school vocational education courses because of the current period utility she derives, because of its effects on her future PSE institution utility, or because of its effects on her future wages in each occupation. By identifying these present and future effects of each choice, the parameter estimates of the dynamic discrete choice model provide more detail about the relationship between the explanatory and dependent variables and more context about what drives

[17] The parameter relating wage utility to non-pecuniary utility (φ) is identified because wages are observed and the distribution of the non-pecuniary utility error terms is assumed to be EV(0,1). As wage and non-pecuniary utility parameters are identified as discussed above, the relationship between wage and non-pecuniary utility is identified by co-variation in observed wages and individual choices each period.

individual decision making.

Second, with a structural model I can jointly estimate effects that pertain to multiple, interrelated research questions. For example, by estimating a structural model, I can jointly estimate the effects of high school vocational education on wages in each occupation, the likelihood of being employed in a skilled occupation, the likelihood of graduating from high school, and the likelihood of graduating from a PSE institution, as opposed to estimating each of these effects separately.

Third, I can use the structural model to conduct policy simulations. It is worth noting that some policy simulations can be conducted using non-structural models. For example, the effects of increasing vocational high school opportunities nationwide could be simulated by adding vocational high school opportunities in the first stage of a 2SLS regression for every individual in the data set and seeing how the addition of these opportunities, for the subset of the sample that did not previously have access to them, would affect predicted values for aggregate wages and employment outcomes. For this simulation, the main benefit of the structural estimation approach is improved sample fit caused by accounting for forward-looking behavior and applying structure to the model (for examples of the general model fit and out-of-sample fit benefits provided by structural models, see Todd & Wolpin's (2006) model of Progressa, Duflo, Hanna & Ryan's (2011) model of teacher attendance decisions in India, and Kaboski & Townsend's (2011) model of microfinance programs in Thailand).

However, many policy simulations cannot be conducted without a structural model of forward-looking behavior. This class of simulations includes policy simulations that force individuals down alternative choice paths, those that change the structure of the model in a substantive ways, and those that change the intertemporal effects of different choices (such as how decreasing the cost of community college would effect an individual's high school decisions). By estimating a structural model I can simulate the effects of these types of policies and predict how they would affect an individual's decisions throughout her lifetime.

F. Variable Construction

F.1 High School Curricula Construction Rules

High School Course Mapping

The transcript courses in ELS:2002 are coded using the Classification of Secondary School Courses (CSSC), a coding system based on the High School Transcript Studies conducted by the NCES (National Center for Education Statistics, 2000). All U.S. high school courses are coded with a six digit code, organized by course type. The first two digits, which denote the main program area, range from 01 – 56. See Table F2 for how these codes are mapped to the five high school fields in my model (*Academic, General Education, Business Vocational, Trade Vocational, and Other Curriculum*).

This mapping schema roughly follows the mapping used by Meer (2007), with the exception that I have added a fifth category, "other", which Meer instead spread across the general education, trade vocational, and business vocational fields. I separate "other" courses to restrict them from impacting the parameter estimates associated with general education, trade vocational, and business vocational high school curricula.

Table F1: CSSC Code Mapping

Course Content	CSSC Code
Academic Courses	
Area and Ethnic Studies (Honors)	050105, 050116, 050120, 050126
Computer and Information Sciences (Honors/AP/IB)	110132-44, 110212, 110213
Engineering	14****
Foreign Languages (Honors/AP/IB/CEEB Prep)	160517, 160544, 160545, 160556, 160907, 160917, 160937, 160943-52
Letters/English (Honors/AP/IB)	230102, 230105, 230108, 230111, 230114, 230117, 230165-71
Liberal/General Studies (Gifted / College Level)	240141, 240151
Life Sciences (Honors/AP/IB)	260141-46
Mathematics (Honors/AP/IB/Advanced)	270410, 270414, 270415, 270417-20, 270424, 270429-35, 270532
Multi/Interdisciplinary Studies (IB/Advanced)	300112-21, 300623
Philosophy and Religion (IB)	380142
Physical Sciences (Honors/AP/IB/ Advanced)	400300, 400521-41, 400622, 400821-31
Psychology (AP/IB)	420114, 420115
Social Sciences (Honors/AP/IB)	450613-16, 450711, 450803, 450806, 450808, 450836, 450850, 450853, 450856, 450870-74, 450921, 451013, 451015, 451018, 451034-37, 451171-81
General Education Courses	
Area and Ethnic Studies (non-honors)	05****
Foreign Languages (non-honors)	16****
Letters/English (non-honors)	23****
Liberal/General Studies (non-honors)	24****
Life Sciences (non-honors)	26****
Mathematics (non-honors)	27****
Multi/Interdisciplinary Studies (non-honors)	30****
Philosophy and Religion (non-honors)	38****
Physical Sciences (non-honors)	40****
Science Technologies	41****
Psychology (non-honors)	42****
Public Affairs	44****
Social Sciences (non-honors)	45****
Business Vocational Courses	
Business and Management	06****
Business and Office	07****
Marketing and Distribution	08****
Communications (except Journalism and Special languages)	09****
Computer and Information Sciences (non-honors)	11****
CTE Business and Office	552***

Trade Vocational Courses	
Communications Technologies	10****
Consumer, Personal, and Miscellaneous Services	12****
Engineering and Engineering-related Technologies	15****
Industrial Arts	21****
Protective Services	43****
Construction Trades	46****
Mechanics and Repairers	47****
Precision Production	48****
Transportation and Material Moving	49****
CTE Industrial Arts, CTE Precision Production, CTE Trades & Industrial Construction, CTE Mechanics & Repairers, Service Occupations	555***, 557***, 558***, 559***

Other Curriculum Courses	
Architecture and Environmental Design	04****
Communications (Journalism and Special languages)	0904**, 0908**
Education	13****
Home Economics	19****
Vocational Home Economics	20****
Law	22****
Summer Abroad, Independent Study, Other Liberal/General Studies	240121, 2401131, 240100
Library and Archival Sciences	25****
Military Sciences	28****
Military Technologies	29****
Parks and Recreation	31****
Citizenship/Activities	33****
Health Related Activities	34****
Interpersonal Skills	35****
Leisure and Recreational Activities	36****
Personal Awareness	37****
Theology	39****
Visual and Performing Arts	50****
Executive Internship	51****
General EMH (Including Pre-vocational Programs)	52****
Special Education	54****
Vocational Career Prep / Exploration, CTE Home Economics	550***, 554***
Special Education – Resource Curriculum	56****
Agribusiness and Agricultural Production	01****
Agricultural Sciences	02****
Renewable Natural Resources	03****
CTE Agriculture	551***
Allied Health	17****
Health Sciences	18****
CTE Health Occupations	553***
Basic Skills	32****

Notes:

1) "*" Indicates that all courses within the program area, not listed elsewhere, fall within the stated course content.

Yearly Curriculum Construction Rule Details

I assign a yearly field concentration to each year of high school based on the credit hours and field types of the classes the individual passed during the year. Each individual takes six credit hours of classes in a given year.[18] I assign yearly field concentration as described below. This specification is similar to other specifications used in the literature, such as Meer (2007).

- The year is coded as a Trade Vocational yearly field concentration if the individual took more Trade Vocational credits than either Business Vocational credits or Academic credits *AND* took 1.25 or more Trade Vocational credits.
- The year is coded as a Business Vocational yearly field concentration if the individual took more Business Vocational credits than either Trade Vocational credits or Academic credits *AND* took 1.25 or more Business Vocational credits.
- The year is coded as an Academic yearly field concentration if the individual took more Academic credits than either Trade Vocational credits or Business Vocational credits *AND* took 1.25 or more Academic credits.
- The year is coded as a General Education yearly field concentration if the individual took 1.25 or more General Education credits *AND* took less than 1.25 Trade Vocational credits, took less than 1.25 Business Vocational credits, took less than 1.25 Academic credits, and took less than 2 Other Curriculum credits.
- The year is coded as an Other Curriculum yearly field concentration if the individual took 2 or more Other Curriculum credits *AND* took less than 1.25 Trade Vocational credits, took less than 1.25 Business Vocational credits, and took less than 1.25 Academic credits.
- The year is coded as an Other Curriculum yearly field concentration if an individual took less than 1.25 credits in each of the other four fields.
- In the event of ties, the tiebreaking order is Trade Vocational, Business Vocational, Academic.[19]

Alternative Curriculum Construction Rules

[18] Credit hours from schools that assign a different number of credit hours in a year (e.g. 12 credit hours per year) are first adjusted so that the average number of credit hours taken by a full time student at that school each year is six.
[19] 0.2% of student-year curricula observations had ties. Using alternative tiebreaking orders does not affect the 2SLS estimation results.

I investigated three alternative curriculum construction rules. The first rule defines an individual's overall curriculum as the yearly field concentration (constructed as described above) taken during her senior year. The second rule aggregates a student's classes and credit hours across all four years of high school and then chooses an overall concentration based on aggregate credit hours in each field.[20] Finally, the third rule defines an individual's overall curriculum as the value of the pre-constructed variable in the ELS:2002 data set that assigned high school graduates to either an academic, occupational, academic & occupational, or other curriculum.

See Table F2 for a comparison of how aggregate outcomes change with each of the four construction rules. The table shows that curriculum outcomes are very similar across all four construction rules.

Table F2: Curriculum Construction Rule Comparison

HS Curriculum	Constructed Outcomes	Alternative 1: Senior Year Classes	Alternative 2: All Classes	Alternative 3: ELS Concentrations	
Academic	27.2%	25.2%	30.6%	*Academic*	24.6%
Gen Ed	42.9%	46.0%	44.3%	*Occupational*	12.6%
Bus Voc	7.0%	6.7%	5.8%	*Acad & Occ*	2.3%
Trade Voc	5.8%	5.2%	5.8%	*Other*	60.4%
Other	17.2%	16.8%	13.5%		

Notes:
Total # of classifiable observations varies across construction rules based on available data, from 11,880 to 14,810. Sample sizes are rounded to the nearest ten to comply with secure data disclosure requirements.

F.2 Employment Construction Rules

An ELS:2002 survey participant denoted her occupations between the years of 2002 and 2012 using six-digit O*NET occupation codes. ELS:2002 survey staff then mapped these six-digit O*NET occupation codes to one of 14 constructed occupations (Ingels et al., 2014). See Table F3 for how these 14 occupations are mapped into the five occupation choices in my model (*Professional, Skilled Manuel Labor, Skilled Non-manual Labor, Skilled Other, and Unskilled*).

[20] For this alternative construction rule I followed the yearly field concentration rules as defined above, except with slightly different credit assignment ratios (taking the place of 1.25 out of 6 and 2 out of 6): 3 out of 24 for trade vocational, business vocational, and academic, 6 out of 24 for general education, and 8 out of 24 for other curricula. These ratio's were chosen to take into account the large number of general education and other curricula courses that individual's take during their first and second years of high school, and to roughly follow the construction rules used in the previous literate (e.g. Meer, 2007).

Note that, while the other 12 occupation codes provided in ELS:2002 fit directly into one of the five employment categories, the laborer and service occupations do not as they aggregate both skilled and unskilled workers together. As such, to construct the unskilled occupation I further split these employment categories between the skilled manual labor, skilled non-manual labor, and unskilled occupations based on the 6-digit O*NET occupation code provided in the data set for each occupation.

Table F3: 2002-2012 Occupation Code Mapping

Coded Occupation	ELS:2002 Occupation
Professional	
	Manager, Administrator
	Professional A
	Professional B
Skilled Manual Labor	
	Craftsperson
	Operative
	Technical
	Protective Service
	Laborer (skilled, see notes)
Skilled Non-Manual Labor	
	Clerical
	Sales
	Service (skilled, see notes)
Skilled Other	
	Farmer, Farm Manager
	Military
	School Teacher
Unskilled	
	Laborer (unskilled, see notes)
	Service (unskilled, see notes)

Notes:
Based on six-digit O*NET codes, the following Laborer and Service occupations were coded as Unskilled Occupations: Merchandise Displayers and Window Trimmers; Lifeguards, ski patrol, and other recreational protective service workers; cooks – fast food; food prep, bartenders, counter attendants, waiters, hosts, dishwashers; janitors and cleaners; Attendants (service stations, ticket takers, etc); bellhops; and cashiers. All other Laborer and Service occupations were coded as Skilled Manual Labor and Skilled Non-Manual Labor occupations, respectively.

An ELS:2002 survey participant denoted her occupations prior to 2002 by selecting one of 15 occupation types. The 15 occupation types available were chosen by ELS:2002 survey staff. See Table F4 for how these 15 occupations are mapped into the five occupation choices in the model (*Professional, Skilled Manuel Labor, Skilled Non-manual Labor, Skilled Other, and Unskilled*).

Table F4: 2000-2001 Occupation Code Mapping

Coded Occupation	ELS:2002 Occupation
Professional	
	No Codes
Skilled Manual Labor	
	Construction work
	Beautician, hair stylist, barber
Skilled Non-Manual Labor	
	Salesperson, customer service
	Computer related job
	General office or clerical worker
Skilled Other	
	Farm worker
	Hospital or health worker
Unskilled	
	Food service/server/host/dishwasher
	Babysitter or child care
	Cashier, grocery clerk/bagger
	Lawn work or odd jobs
	Camp counselor/lifeguard/coach
	Warehouse worker
	House cleaning or janitorial work
Unknown Occupation	
	Other

F.3 Local Labor Market Characteristic Construction Rules

Local Labor Market Industry Mapping

Bureau of Economic Analysis Local Area Personal Income & Employment data (U.S. Bureau of Economic Analysis, 2002) contains county-level employment percentages for each two-digit NAICS industry. See Table F5 for how each two-digit NAICS industries is mapped into one of four constructed industries (*Professional Industries, Skilled Manuel Labor Industries, Skilled Non-manual Labor Industries, and Other Industries*).

Table F5: Local Labor Market Industry Mapping

Industry	NAICS Industry Code
Professional	
Professional, Scientific, and Technical Services	54
Management of Companies and Enterprises	55
Skilled Manual Labor	
Mining	21
Utilities	22
Construction	23
Manufacturing	31-33
Transportation and Warehousing	48-49
Waste Management	562
Other Services (Repair and Maintenance)	811
Skilled Non-Manual Labor	
Wholesale Trade	42
Retail Trade	44-45
Information	51
Finance and Insurance	52
Real Estate and Rental and Leasing	53
Other	
Farm Employment	NA
Agriculture, Forestry, Fishing, and Hunting	11
Administration	561
Educational Services	61
Health Care and Social Assistance	62
Arts, Entertainment, and Recreation	71
Accommodation and Food Services	72
Other Services (Everything except Repair and Maintenance)	812,813,814
Public Administration	92

Industry / Occupation Comparison

As noted in Section 4.1, local labor market industry variables are used because local labor market occupation data is not available at the county level. However, occupation data is only available at the national level and for each metropolitan statistical area in the United States. This occupation data is available from the Bureau of Labor Statistics' Occupational Employment Statistics (OES) program (U.S. Bureau of Labor Statistics, 2002). In order to compare industry percentages and occupation percentages at the national and MSA level, I first map each Standard Occupation Classification (SOC) System occupation code into one of four constructed occupations (*Professional Occupations, Skilled Manuel Labor Occupations, Skilled Non-manual Labor Occupations, and Other Occupations*) as described in Table F6.

Table F6: Local Labor Market Occupation Mapping

Occupation	SOC Code
Professional	
Management Occupations	11
Computer and Mathematical Occupations	15
Architecture and Engineering Occupations	17
Life, Physical, and Social Science Occupations	19
Legal Occupations	23
Healthcare Practitioners and Technical Occupations	29
Skilled Manual Labor	
Construction Trades and Extraction Workers	47
Installation, Maintenance, and Repair Workers	49
Production Occupations	51
Transportation and Material Moving Occupations	53
Skilled Non-Manual Labor	
Business Operations and Financial Specialists	13
Sales Occupations	41
Office and Administrative Support Occupations	43
Other	
Community and Social Science Occupations	21
Education, Training, and Library Occupations	25
Arts, Design, Entertainment, Sports, and Media Occupations	27
Healthcare Support Occupations	31
Protective Service Occupations	33
Food Preparation and Serving Occupations	35
Building and Ground Cleaning and Maintenance Occupations	37
Personal Care and Service Occupations	39
Farming, Fishing, and Forestry Occupations	45
Military Specific Occupations	55

Table F7 presents a comparison of national occupation percentages and industry percentages using my constructed occupations and industries. Table F8 provides the average difference across MSAs between occupation percentages and industry percentages. Finally, Table F9 provides a more detailed crosswalk between national occupation percentages, broken down by OCCSOC codes, and national industry percentages, broken down by two-digit NAICS industry codes. Tables F7, F8, and F9 show that the industry mapping used to create my local labor market characteristic variables reasonably reflect OCC occupation mappings at the national level and within MSAs. As such, my local labor market industry percentages are likely a good proxy for local labor market occupation percentages at the county level.

Table F7: National Constructed Industy / Occupation Crosswalk

Occupations	*Industries*			
	Professional	Sk. Manual Labor	Sk. Non-Manual Labor	Other
Professional	**47.5%**	10.1%	11.6%	16.8%
Sk. Manual Labor	5.8%	**71.7%**	16.0%	8.7%
Sk. Non-Manual Labor	38.8%	16.0%	**65.0%**	20.7%
Other	7.8%	1.9%	6.7%	**53.7%**

Notes:
Constructed occupations are on the y-axis, and constructed industries are on the x-axis. Data is from 2002.

Table F8: Constructed Industy / Occupation Variable Difference Across MSAs

	Professional	Sk. Manual Labor	Sk. Non-Manual Labor	Other
Mean	-9%	-1%	-8%	17%
Std Dev	3%	5%	3%	6%

Notes:
Percentages are industry percentages minus occupation percentages. Data is from a comparison of 295 MSAs in 2002.

Table F9: Detailed National Crosswalk

SOC Occupations	*NAICS Industries* Professional		Sk. Manual Labor										Sk. Non-Manual Labor						Other							
	54 - Professional Services	55 - Management	21 - Mining	22 - Utilities	23 - Construction	31 - Manufacturing A	32 - Manufacturing B	33 - Manufacturing C	48 - Transportation A	49 - Transportation B	562 - Waste Management	811 - Repair	42 - Wholesale Trade	44 - Retail Trade A	45 - Retail Trade B	51 - Information	52 - Finance & Insurance	53 - Real Estate & Leasing	11 - Agriculture & Forestry	561 - Administration	61 - Education	62 - Health Care	71 - Arts	72 - Lodging & Food	812-4 - Other Services	92 - Public Administration
Professional																										
11 - Management	9%	16%	6%	6%	6%	4%	6%	6%	4%	4%	6%	4%	7%	4%	3%	8%	9%	11%	3%	3%	5%	4%	5%	4%	8%	6%
15 - Computer / Mathematical	11%	9%	1%	3%	0%	0%	1%	2%	0%	1%	0%	0%	3%	1%	0%	10%	5%	1%	0%	1%	1%	0%	0%	0%	1%	2%
17 - Architecture / Engineering	16%	3%	4%	8%	1%	1%	2%	8%	0%	1%	2%	0%	1%	0%	0%	2%	0%	0%	0%	1%	0%	0%	0%	0%	0%	3%
19 - Sciences	5%	2%	3%	2%	0%	1%	3%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	1%	0%	1%	1%	0%	0%	0%	3%
23 - Legal	6%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	3%
29 - Health Technical	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	1%	0%	2%	2%	32%	1%	0%	0%	4%
Sk. Manual Labor																										
47 - Construction / Extraction	1%	1%	33%	7%	67%	0%	2%	2%	1%	0%	18%	1%	0%	0%	0%	0%	0%	2%	0%	2%	0%	0%	0%	0%	0%	5%
49 - Maintenance / Repair	1%	2%	9%	26%	8%	5%	6%	4%	8%	3%	6%	50%	7%	7%	2%	8%	0%	14%	2%	2%	1%	1%	3%	1%	1%	4%
51 - Production	2%	2%	10%	12%	1%	55%	49%	54%	1%	3%	3%	7%	7%	3%	1%	4%	0%	1%	4%	8%	0%	1%	0%	1%	9%	2%
53 - Transportation	1%	3%	17%	2%	3%	16%	13%	5%	56%	58%	44%	18%	20%	9%	6%	3%	0%	7%	13%	10%	3%	1%	2%	2%	5%	4%
Sk. Non-Manual Labor																										
13 - Business Ops / Financial	9%	14%	4%	7%	2%	1%	2%	3%	2%	2%	2%	1%	3%	1%	1%	4%	19%	3%	0%	2%	2%	1%	1%	0%	8%	8%
41 - Sales	4%	5%	1%	2%	2%	4%	3%	2%	2%	3%	2%	5%	25%	50%	56%	13%	13%	22%	1%	9%	0%	0%	8%	4%	6%	1%
43 - Office & Admin Support	24%	34%	10%	23%	9%	8%	11%	10%	19%	25%	12%	11%	24%	14%	20%	23%	50%	22%	5%	22%	11%	17%	10%	4%	17%	25%
Other																										
21 - Community / Social Sci	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	6%	0%	0%	4%	4%
25 - Education	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	58%	3%	2%	0%	3%	2%
27 - Arts	4%	1%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	2%	16%	0%	1%	0%	0%	1%	0%	9%	0%	2%	1%
31 - Heath Support	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	19%	0%	0%	1%	2%
33 - Protective Service	0%	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	2%	0%	13%	1%	0%	4%	1%	1%	18%
35 - Food Prep & Serving	0%	1%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	4%	1%	2%	0%	2%	0%	1%	4%	3%	18%	75%	4%	1%
37 - Building Maintenace	1%	1%	0%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%	1%	0%	1%	0%	10%	1%	21%	5%	3%	10%	6%	4%	2%
39 - Personal Care	1%	1%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	3%	0%	2%	1%	1%	2%	7%	26%	2%	24%	3%
45 - Farming & Forestry	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	68%	0%	0%	0%	0%	0%	0%	0%
55 - Military	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Notes:

Percentages aggregate vertically. They display the occupational breakdown of types of job in each industry nationwide. Constructed occupations are on the y-axis, and constructed industries are on the x-axis. Data is from 2002.

F.4 Additional Variable Construction Details

ELS:2002 Raw Variables

This subsection provides additional information about the data elements available in ELS:2002 that are used to construct the variables used in this paper. See table F10 for a list of the variables in the ELS:2002 raw data file that are used to create log-hourly wage, high school attendance, PSE attendance, and employment outcomes each year. The Stata do files that map this information into the variables discussed in Section 4 are available upon request.

Note that most wages in ELS:2002 were collected as hourly wages, although for a subset of student-year observations weekly, monthly, or yearly income was collected instead. These incomes are first converted to hourly wages based on the number of hours each individual worked per week and the number of months they worked throughout the year. Also note that wages are adjusted for inflation into 2002 dollars. Hourly wages below 5 dollars an hour and above 100 dollars an hour are dropped (nine percent of hourly wages are dropped because they were below $5 an hour, and one half of one percent of hourly wages are dropped because they were above $100 an hour).

Additional Imputation Rules

As shown in Table A6, choice information is missing for many student-year observations in the data set. In addition, conflicting choice information is provided for a small number of student-year observations. See Table F11 for details on how a subset of these missing student-year observations are imputed based on available data as well as how conflicting choice information is coded. The Stata do files that contain these rules are available upon request.

Table F10: Raw Variables

Constructed Variable / Raw Variables
High School Attendance
High School Transcript data 2000-2003 (courses, credits, grades)
Year/month graduated from high school, and type of graduation (GED or diploma)
Year/month left high school (prior to 2006) and why (graduated, dropped out, or transferred)
High School grade attended in 2003
Enrolled in the spring of 2003 (Y/N)
Enrolled in the spring of 2004 (Y/N)
Working towards graduation (GED) in 2012 (Y/N)
Failed 9th or 10th grade (Y/N)
Failed 11th or 12th grade (Y/N)
Dropouts
Year/month first dropped out
Year/month first returned
Year/month second dropped out
Last grade attended before dropping out and whether passed/failed
Attended High School in 2002 (Y/N)
Post-Seconardry Education Attendance
Year/month first began attending a PSE intuition, and institution type
Year/month began attending most recent PSE intuition, and institution type
Year/month last attended most recent PSE intuition, and institution type
Year/month first received a PSE degree, and degree type
Year/month received highest PSE degree, anddegree type
Ever attended a PSE (asked in both 2006 and 2012) (Y/N)
Attended a PSE institution, and institution type, monthly from 2003 to 2005 (Y/N)
Attending a PSE institution in 2012 and institution type (Y/N)
Employment
Prior to Jun 2012: Occupation type and year/month began and ended most recent job
Prior to Jan 2006: Occupation type and year/month began and ended first job after high school
Occupation type and year/month began the job employed in during Jan 2006.
Prior to May 2002: Occupation type and year/month began and ended most recent job
Occupation and hours worked a week in 2001
Whether working in 2012
Number of weeks employed in 2011
Whether working for six or more months in 2010 and 2009
Whether employed each month from Jun 2002 to Jan 2006
Number of hours worked a week in 2001 and 2003
Whether working in 2003
Year/month began and ended most recent job (as of '03), only for dropouts and early graduators
Log Hourly Wages
Wages current / most recent job as of 2012
Wages in 2011
Wages in 2005
Wages first job after school (prior to Jan 2006)
Wages current / most recent job (as of Jan 2006)
Wages current / most recent job (as of 2003), only for dropouts and early graduators

Table F11: Additional Interpolation Rules (for Years with Missing Data)

Constructed Variable / Raw Variables
High School Attendance
On Time Graduates : Individuals that graduated on time (in 2003) are coded as attending in 2000, 2001, and 2002
Early Graduates : Individuals that graduated early (in 2002 or 2001) are coded as attending in 2000 and 2001, and are coded as already having finished 1-2 years of high school (respectively) prior to 2000
Late Graduates : Individuals that graduated after 2003 are coded as attending in the year of graduation
All years after graduation are coded as not attending
Dropouts : Every year after final dropout, including final dropout year, is coded as not attending. Every year before first dropout year is coded as attending. If dropped out twice, year of return is coded as attending and year of first dropout is coded as not attending
Post-Seconardry Education Attendance
Individuals that attended a PSE institution for at least six months in a year are coded as attending that year
Every year before began attending first PSE institution is coded as not attending
Every year after last began attending most recent PSE institution is coded as not attending
If the first year attended and most recent year attended are both at 4-yr institutions, and the years are four years apart, the two years between them are coded as attending 4-yr institutions
Employment
Code as working (type unknown) if worked more than six months each year
Code as working if worked more than 20 hours a week in 2001
Conflicting Data
When data on whether an individual attended HS / PSE full-time or part-time is missing, I code individuals as follows:
a. Individuals that attended school (credit amount unknown) and worked during any year between 2000 and 2008 are coded as attended school
b. Individuals that attended school (credit amount unknown) and worked during any year between 2009 and 2012 are coded as worked

Notes:
These interpolation rules are only used for student-years for which outcomes are unobserved in the data.

References

Berndt, Ernst, Bronwyn Hall, Robert Hall, and Jerry Hausman, 1974. "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement* 3: 653–665.

Biden, Joe, 2019. "The Biden Plan for Education Beyond High School." *https://joebiden.com/beyondhs/*, retrieved Jan 26, 2020.

Bishop, John H. and Ferran Mane, 2004. "The Impacts of Career-Technical Education on High School Labor Market Success." *Economics of Education Review* 23: 381-402.

Borsch-Supan, Axel and Vassilis Hajivassiliou, 1993. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Depended Variable Models." *Journal of Econometrics* 58: 347-368.

Castellano, Marisa, James R. Stone III, and Sam Stringfield, 2005. "Earning Industry Recognized Credentials in High School: Exploring Research and Policy Issues." *Journal of Career and Technical Education* 21(2): 7-34.

Chan, Marc K., 2013. "A Dynamic Model of Welfare Reform." *Econometrica* 81: 941-1001.

Clinton, Hillary, 2016. "Making college debt-free and taking on student debt." *https://www.hillaryclint on.com/issues/college/*, retrieved October 13, 2016.

Duflo, Esther, Rema Hanna, and Stephen Ryan, 2011. "Incentives Work: Getting Teachers to Come to School," American Economic Review 102(4): 1241-1278.

Geweke, John, 1988. "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference." *Journal of Econometrics* 38, 73-89.

Hossler, Don, Doug Shapiro, Afet Dundar, Mary Ziskin, Jin Chen, Desiree Zerquera, and Vasti Torres, 2012. "Transfer & Mobility: A National View of Predegree Student Movement in Postsecondary Institutions." National Student Clearinghouse Research Center and Project on Academic Success, Indiana University.

Ingels, Steven J., Daniel J. Pratt, Christopher P. Alexander, Donna M. Jewell, Erich Lauff, Tiffany L. Mattox, and David Wilson, 2014. "Education Longitudinal Study of 2002 Third Follow-up Data File Documentation." (NCES 2014-364). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Kaboski, Joseph and Robert M. Townsend, 2011. "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative." Econometrica 79(5): 1357-1406.

Keane, Michael P. and Kenneth I. Wolpin, 1997. "The Career Decisions of Young Men." *The Journal of Political Economy* 105(3): 473-522.

Klein, Steven, Amanda Richards Sheil, Robin White, Sandra Staklis, Corinne Alfeld, Caitlin Rose Dailey, Ivan Charner, and Anne Poliakoff, 2014. "Evaluation of the Implementation of the 'Carl D. Perkins Career and Technical Education Act of 2006 - Finance, Accountability, and Programs of Study." RTI International, https://eric.ed.gov/?id=ED613352, retrieved February 10, 2022.

Meer, Jonathan, 2007. "Evidence on the Returns to Secondary Vocational Education." *Economics of Education Review* 26: 559-573.

National Center for Education Statistics, U.S. Department of Education, 2016. "High School

Transcript Studies - CSSC Courses/Course Codes." *https://nces.ed.gov/surveys/hst/courses.asp*, retrieved March 4, 2016.

Obama, Barack, 2015. “Remarks by the President in State of the Union Address, January 20, 2015.” Office of the Press Secretary, The White House. *https://www.whitehouse.gov/the-press-office/2015/01/20/remarks-president-state-union-address-january-20-2015*, retrieved September 4th, 2015.

Roza, Marguerite, 2009. “Breaking Down School Budgets: Following the Dollars into the Classroom.” *Education Next* 9(3): 28+.

Sanders, Bernie, 2016. “It’s Time to Make College Tuition Free and Debt Free.” *https://berniesanders.com/issues/its-time-to-make-college-tuition-free-and-debt-free/*, retrieved Oct 13, 2016.

Todd, Petra and Kenneth I. Wolpin, 2006. “Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Behavioral Model of Child Schooling and Fertility.” American Economic Review 96(5): 1384-1417.

Train, Kenneth E., 2003. “*Discrete Choice Methods with Simulation*,” New York, NY: Cambridge University Press.

U.S. Bureau of Economic Analysis, 2002. “CA25N Total Full-Time and Part-Time Employment by NAICS Industry.” *https://www.bea.gov/iTable/iTable.cfm?reqid=70*, retrieved May 9, 2016.

U.S. Bureau of Labor Statistics, 2002. “Occupational Employment Statistics Data, 2002.” *https://www.bls.gov/oes/tables.htm*, retrieved May 9, 2016.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, 2013. “National Assessment of Career and Technical Education: Interim Report.” Washington, D.C.