EXPANSION OF HIGHER EDUCATION AND LABOR MARKET OUTCOMES: EVIDENCE FROM TURKEY

ASENA CANER, TOBB UNIVERSITY OF ECONOMICS AND TECHNOLOGY
MERVE DEMIREL, BILKENT UNIVERSITY
ÇAĞLA ÖKTEN, BILKENT UNIVERSITY

Abstract. In this paper, we estimate the impact of the dramatic increase in the number of universities and available slots in Turkey on labor market outcomes. To identify returns to higher education, we introduce a measure called "intensity of expansion policy" which captures supply side increase in higher education and affects the labor market outcomes of individuals only through higher education. We estimate the returns to higher education by using intensity of policy as an instrument and show that the causal effect of higher education on hourly wages is positive and statistically significant in general (with a lower returns for female compliers).

1. Introduction

This study quantifies the effects of higher education expansion on labor market outcomes and identifies the distribution of the returns to higher education by exploring a natural experiment: higher education expansion in Turkey as the number of universities increased from 72 in 2005 to 129 by 2009 and the number of available slots increased by more than 60%.

The higher education expansion was exogenous and unanticipated from the perspective of university candidates. Prior to the expansion, demand for the higher education was much higher than available slots (Measuring, Selection and Placement Center Statistics (OSYM), 2004) and a centralized competitive examination was and still is applied to ration excess demand for higher education. Hence the first order effects of the expansion are expected to be on the supply side of the education market. Furthermore, the higher education expansion was politically motivated driven by requests from members of the parliament (Arap, 2010). The Turkish general election was to be held in 2007 to elect 550 members of parliament. Prior to this, in 2005, 25 new public universities were planned to open. Due to pressures from members of the parliament who represent different provinces and seek reelection, 16 more public universities were opened than initially planned by 2008. Finally, the expansion occurred in a short period of time and did not evolve endogenously with the changing patterns in demand across time.

The effects of higher education are investigated by using the 2014-2017 Household Labor Force Survey (HLFS), an annually conducted survey by TURKSTAT. The data for available slots in four-year higher education programs are published yearly per each university and field of major by OSYM. The data used in this study is collected from these publications and converted into electronic format. Hence our dataset, includes

Date: July 7, 2018.

information on the number of available slots for each university, field of study, province and year for the 2002-2012 period. We define the intensity of the expansion in higher education as the number of available slots divided per high school in a given region and year.

Whether and to what extent additional educational opportunities do increase an individuals wages and employment prospects are important questions both in the literature and for the policy-makers. Such an inquiry is especially relevant for emerging economies that aim to increase the educational level of their population through substantial changes in their educational systems. We analyze the causal effect of higher education on labor market outcomes, labor force participation, employment and wages by using intensity of reform variable as an instrument in an instrumental variable estimation.

This is the first paper to our knowledge to identify the returns to higher education by exploring the recent educational reforms in Turkey, and the first to quantify the influence of the expansion of the higher education system, caused by these reforms.

In recent years, many countries, including Italy, France, England, Russia and Turkey, have implemented reforms aimed at promoting higher attainment to and increasing equal opportunities in higher education. Whether and to what extent educational expansion does decrease gender gap and regional disparities in education, increase an individual's wages and employment prospects, as well as generate a positive economic return for a country are important questions both in the literature and for policy-makers.

Bratti, Checci and De Blasio (2008) analyzed the effect of higher education supply increases in Italy between 1995 and 1998 on equalizing opportunities in education. They found that the drop-out rates increased parallel to an increase in enrollment rates so the expansion has no effect on education attainment. Oppesidano (2011) found similar results by using the same data, and concluded that reform did not diminish regional differences. Maurin and McNally(2008) have shown that the provision of higher education in France, which is facilitated temporarily, increased labor force participation and wages of those benefitted from this policy. In UK according to Walker and Zhu (2008), the higher education expansion between 1994 and 2006 has no significant effect on the college premium for men, while there is an insignificant rise for women. Also, Kyui (2016) investigates the impact of higher education supply in Russia between 1990 and 2005 on education and labor market outcomes. She has shown that a long with an increase in education levels, college graduates has also gained additional returns in addition to an increase in probability of having higher education. There are several other studies focuses on the impacts of expansion like Blanden, Machin (2004); Gurgond Maurin (2007), Wang, Fleisher,Li,Li (2014).

There are two studies which focus on the 2006-2008 higher education expansion in Turkey: Yılmaz (2014), examines the effects of the increase in number of universities and available slots on narrowing gender gap in college degree at the city level and finds that expansion had a positive effect in reducing the gender gap. Polat (2017) analyzes the effects of expansion on education outcomes by comparing higher education attainment of 18-25 year olds in HLFS 2004 and HLFS 2012. He finds that the effect of family income on educational attainment is significantly lower in 2012 for women in the north and south east of the country.

Our results show that those who attain education as a result of growing access to universities, are significantly more likely to participate in labor market and find a job, compared to the ones with low-education levels. The compliers of expansion policy also earn significantly more than those without a college degree and their earning returns to education is much higher than the average returns in young population.

The remainder of the paper is organized as follows: Section 2 introduces the institutional context and expansion in Turkish higher education system, section 3 describes the data used in the analysis. Section 4, explains the identification strategy and the hypothesis tested in this study. Section 5 estimates the impact of expansion and section 6 estimates returns to education in terms of labor market outcomes by instrumental variable method and the last section states the concluding remarks.

2. Institutional Context

This section introduces the educational system in Turkey and the recent expansion policies which are the focus of this study. Formal education in Turkey consists of pre-primary education, primary education, secondary education and higher education. The primary education includes both four years of elementary and four years of middle school education which are compulsory for all citizens. Secondary education includes a general or vocational or technical high schools which lasts for four years. Secondary education is also compulsory since 2013. All the compulsory education is free at the public schools.

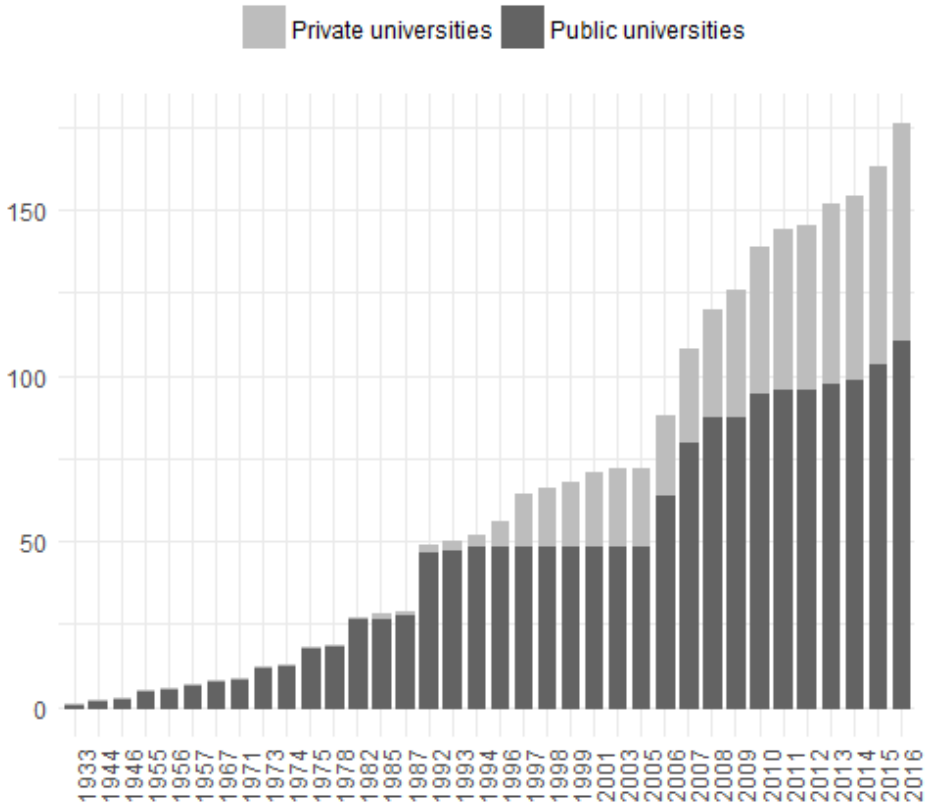
The higher education in Turkey includes all post-secondary education including post-graduate level education. In 1974, application to universities was centralized with the introduction of central university exam and in 1981, with the new Higher Education Law, Council of Higher Education was introduced to regulate higher education. The higher education programs are divided into four according to their levels as vocational school, undergraduate, graduate and doctorate respectively.

There are two types of universities in Turkey namely state and non-profit foundation universities. The first public university was introduced in 1933 and until 1992 there has been a steady increase mostly in the number of public universities. Between 1992 and 2005, the number of public universities remain fairly constant and there is a small increase in the number of non-profit foundation (private) as shown in Figure 1. The total number of universities increased from 72 in 2005 to 129 by 2009 and the number of available slots increased by more than 60%. This expansionary period is the focus of our study.

Prior to the expansion, demand for the higher education was much higher than available slots (2004 Measuring, Selection and Placement Center Statistics) and a centralized competitive examination was and still is applied to ration excess demand for higher education.

The higher education expansion was politically motivated driven by requests from members of the parliament (Arap, 2010). The Turkish general election was to be held in 2007 to elect 550 members of parliament. Prior to this, in 2005, 25 new public universities were planned to open. Due to pressures from members of the parliament who represent different provinces and seek reelection, 16 more public universities were opened than initially planned by 2008.

FIGURE 1



Source: Council of Higher Education, Higher Education Statistics

In 2005 prior to expansion, 43 out of 81 provinces did not have a university whereas by 2009 only 4 remained without a university. The number of provinces with multiple universities increased from 5 to 8. In many of the provinces, the number of available slots in higher education more than doubled. As a result, the number of students who enrolled in a four year program in a higher education institution and schooling rate in the higher education also increased dramatically (Table 1).

The Council of Higher Education (COHE) serves as the head of all higher education institutions since 1981 and the available slots for each university and each academic track is monitored by the COHE. According to 2007 report of COHE, the major increase is aimed at four year undergraduate programs because the Measuring, Selection and Placement Center stated that the demands for vocational schools are not higher than the available slots according to 2004 statistics, while the demand for four year undergraduate programs is much higher than the available slots.

When reform was adopted at the end of 2005, the Council of Higher Education opposed the draft by saying that there was not enough resources and infrastructure for the establishment of these new universities. Indeed, despite the fact that a total of 24 new universities have been established in 2006 and 2007, there has been no increase in the number of students newly entering the universities or available slots in higher

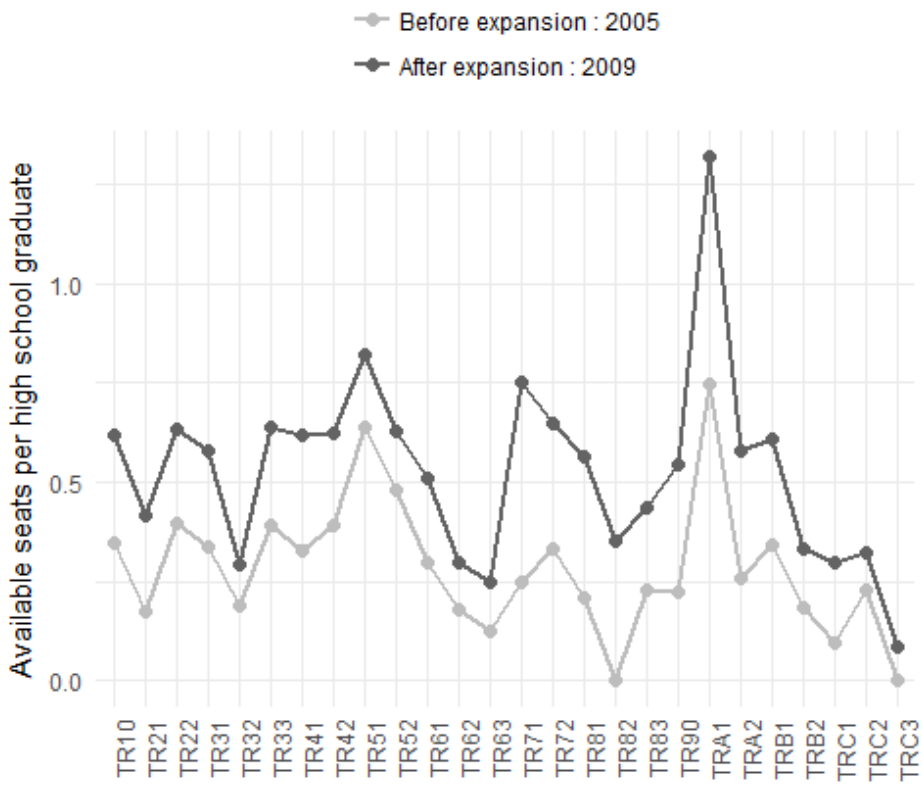
Table 1

	2005	2006	2007	2008	2009	2010
Students who take the exam	1730876	1678383	1776441	1646376	551304	1588624
Quotas for 4-year programs	197296	195910	201463	267502	319128	349472
New students in 4-year programs	194516	177258	193541	264088	300029	343410

Source: Measuring, Secelction and Placement Center Statistics

education programs these years. The Table 1 represents the higher education statistics across years in the country. It shows that the main increase in available slots and new students in higher education is observed in 2008, so in this paper we determine the years of post-expansion as 2008 and after.

Figure 2



Source: Measuring, Selection and Placement Center, Yearly Periodicals, Authors' calculations.

The adjusted number of available slots per high school graduates in each of the 26 NUTS-2 regions is our measure of higher education supply in this study and 2shows the regional variation in this variable in pre-expansionary period 2005 and post expansionary period 2009. The graph clearly shows that, in each period, there is a large regional variation in available slots per high school graduates and there is also a variation in the increase of this variable from 2005 to 2009 across regions. a natural experiment: higher education expansion in Turkey as the number of universities increased from 72 in 2005 to 129 by 2009 and the number of available slots increased by more than 60%.

3. Data Description

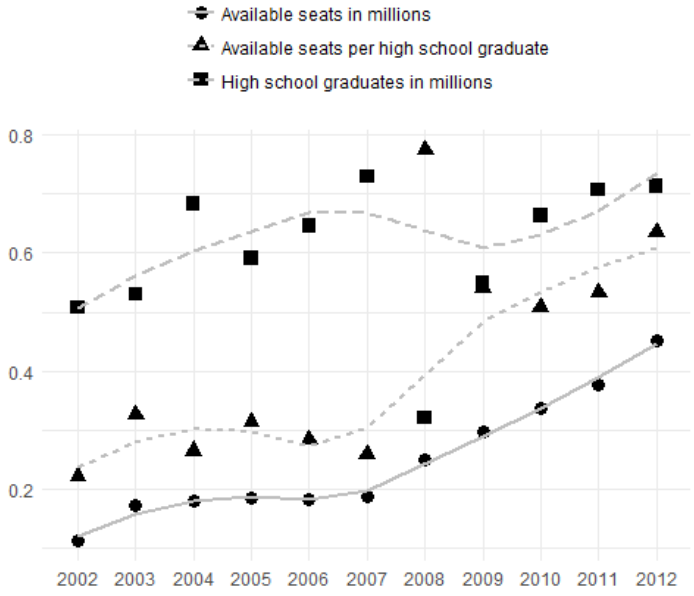
The data set is formed by combining data from three different sources. The first data source is Measuring, Selection and Placement Center (OSYM)'s Periodicals. OSYM publishes a document which includes the available slots of four-year higher education programs (majors) in each university, the number of students who are placed in these programs, and the test score ranges of the settlers in a given year. We collected this pdf data available online and converted it into electronic format. Hence, we constructed a dataset including the available slots in each program and university in a given year for the 2002-2012 period. We also collected region information for each university and merged this region information indicating the province and NUTS-2 level regional identifier of each university to this data set. This data set allows to construct several measures for the supply of higher education opportunities in Turkey.

Our second data source is the number of high school graduates at the city level by TURKSTAT. The high school graduates data represent the population that potentially can benefit from higher education opportunities at the regional level.

We construct a variable called *intensity of policy*, defined as the available slots divided by the number of high school graduates in a given year, by combining the two data sets mentioned above. This variable is constructed both at the region level and country level. he intensity variable measures educational opportunities in a given region and year. The number of high school graduates represents the potential demand for higher education in that year, while the available slots in higher education programs gives the supply of higher education in a given year at the region level. Figure 3 shows the evolution of demand and supply side in higher education and also the intensity of policy measure across years at the country level.

Third data source is 2014-2016 Household Labor Force Surveys, which are surveys carried out annually by TURKSTAT. In the first part of the questionnaire, there is information about households, sampling, personal characteristics of household members (age, gender, marital status, etc.). This section also contains information about the educational status of the household (the highest degree that the individual has taken). The second part of the questionnaire contains information on employment and income. Finally, the location information is given according to the first and second level of Statistical Region Units Classification (NUTS-1 and NUTS-2) which divides Turkey into 12 and 26 statistical regions respectively. Since the most detailed information about the region of residence is available at NUTS-2 level, we have introduced all regional characteristic variables, like available slots per high school graduate at NUTS-2 level and define region dummies at NUTS-1 level to avoid multicollinearity problems.

FIGURE 3



The data from the Household Labor Force Survey between the years 2014-2016 is limited to the individuals expected to have started university education between 2002 and 2012. There is no information on which year the individuals have graduated from high school or take the university entrance exam. Using Measuring, Selection and Placement Center statistics, we assume individuals start higher education at age 18. Thus, we restrict our sample to individuals who were 18-years old in 2002-2012 (simply will be called cohorts of 2002-2012 for the rest of the paper).

The last dataset used in the study is the regional gross domestic product per capita. This data is obtained from TURKSTAT at NUTS-2 level for years 2004-2014. Along with the increase in university quotas in a region, changes in macroeconomic conditions in different areas of development (trade development, new roads, etc.) may have affected the desire to pursue university education. For example, the approach in Duflo (2001) is similar to the control of developments outside the field of education. In this study, domestic output per capita will be used as a sign of macroeconomic conditions.

The surveys report the information about the age, gender, level of educational attainment, field of highest degree, region of residence, hours of work and earnings of individuals, including bonus payments and premiums. The information on educational opportunities from the previous data sets is combined with the survey data.

We construct two different samples as follows: First, we exclude the individuals not in the labor force at the time of survey and continuing their education. In Turkey, average age of graduation from higher education is 22, so we also exclude individuals younger than 23 to decrease sample selection. This is the first sample, called *whole sample* that we use to analyze labor market outcomes like labor force participation, employment etc. Then we further restrict whole sample to full-time (30-72 hours a week) wage earners with observable wages and exclude top and bottom %1 of wage distribution to exclude outlier observations, and we refer this sample as *wage sample*

for the rest of the paper and we use this sample to analyze wage outcomes. In the whole sample, women make up %50.7 of the population, while in the wage sample only %29.1 of the observations is constituted by women.

The education status of an individual is represented by dummy variable *higher education attainment* which takes value 1 for those who graduated from college and 0 for the rest. The higher education attainment is %26.3 and %38.5 in the whole sample and wage sample respectively. The proportion of higher education graduates is higher for males but close to males in whole sample: %26.7 of males and % 26.0 of females attain higher education. Among individuals with observable wages, the college graduates constitutes %30.6 of males and %57.6 of females.

Table 2

	Males		*Females*	
	without HE	with HE	without HE	with HE
Labor force participation	0.903	0.903	0.438	0.751
	(0.296)	(0.296)	(0.496)	(0.432)
Employment	0.813	0.791	0.341	0.592
	(0.390)	(0.407)	(0.474)	(0.491)
Hourly wages	5.780	10.836	5.404	10.645
	(2.427)	(5.172)	(2.299)	(5.239)

We define dummy variable *labor force participation*, which is 1 for those who identify themselves as in the labor force, except for those who participate but employed in a family business without payment. Similarly, *employment* dummy variable is 1 for those who identify themselves as employed, except for those who work in a family business without payment.

We use hourly wages to measure earnings and calculated it by using monthly income and weekly working hours, reported in HLFS (monthly income/(weekly hours*4.3)) in a regular week of primary job. We also adjust hourly wages for inflation across years at the country level. The summary statistics of labor market outcome variables labor force participation, employment and hourly wages across higher education attainment are given in Table 2.

An individual's exposure to expansion is determined by the year that individual is 18 years old and the region of residence. The intensity of expansion policy is measured by the available slots per high school graduate at the regional level for each cohort. In HLFS, there is no information about the region of residence of individuals when they are 18 years old, but the information on how long they have been living in their current cities exists in the survey. Thus, we assume they reside in the same region when they are in the year of survey. We also check the robustness of our results for the sample of individuals who did not migrate since their age 16. These are the individuals who were exposed to expansion in that region, since we know the region where they make their higher education decision.

All empirical specifications include region and cohort fixed effects, and we also control for the number of high school graduates whenever we use the instrument, available slots per high school graduate. After controlling for region and cohort fixed effects, intensity of policy measure provides an exogenous measure of capacity constraints in higher education.

4. Instrument Validity

In this section, we describe the explanatory power of the instrument, variations on access to higher education in Turkey, on explaining higher education attainment decision in several specifications of our samples.

We define the higher education attainment binary variable h_{irt} as equal to 1 if individual is college graduate. An individual's cohort and region of residence jointly determines intensity of policy exposed, denoted by s_{irt}. Thus, the education equation, first-stage of simultaneous equations model, is as follows:

$$h_{irt} = \gamma s_{irt} + X'_{h,i}\beta_h + \epsilon_{irt} \tag{4.1}$$

We test four specifications of education equation, by using following candidates for instrumenting higher education attainment.

(1) The number of available slots per high school graduate at the regional level, denoted by *Available slots per high school graduate regional level*.

(2) The number of available slots in millions at the regional level, denoted by *Available slots regional level*.

(3) The number of available slots per high school graduate at the country level, denoted by *Available slots per high school graduate country level*.

(4) The number of available slots in millions at the country level, denoted by *Available slots country level*.

In all specifications, we control for the number of high school graduates at the regional level and country level, for regional and country level instruments respectively. We also control for age, age square, gender, survey year in all specifications. Since the instrument that we define at the regional level varies across NUTS-2 regions (26 regions) and cohorts, we use NUTS-1 region effects and cohort effects in these specifications to avoid perfectly collinear region and cohort fixed effects. However, for the specifications with country level instruments we use NUTS-2 region fixed effects to be able to level the regional differences across the country.

We report the results for education equation in Table **??**: the marginal effects of changes in available slots on higher education attainment from probit edtimation of higher education attainment, with statistics for the tests on exclusion of instruments from the education equation. The results on exclusion restriction test suggest that the instruments at the regional level better explain variations in educational attainment than country level variables, in general. Since there are large variations across regions, we expect to have this result. Also, even though in wage sample *Available slots regional level* has slightly higher explanatory power, *Available slots per high school graduate regional level* performs better over the whole sample. As a result, *Available slots per*

TABLE 3

	Dependent variable: Higher Education Attainment			
	Whole Sample			
Available slots per high school graduate regional level	0.118***			
	(0.0229)			
Available slots regional level		2.151**		
		(1.010)		
Available slots per high school graduate country level			0.892***	
			(0.283)	
Available slots country level				2.173***
				(0.689)
Observations	154,700	154,700	154,700	154,700
LR-test χ^2	26.63	4.54	9.93	9.93
LR-test, P-value	(0.0000)	(0.0331)	(0.0016)	(0.0016)
	Wage Sample			
Available slots per high school graduate regional level	0.0895***			
	(0.0290)			
Available slots regional level		2.333***		
		(0.736)		
Available slots per high school graduate country level			0.777*	
			(0.445)	
Available slots country level				1.892*
				(1.084)
Observations	62,021	62,021	62,021	62,021
LR-test χ^2	9.55	10.03	3.04	3.04
LR-test, P-value	(0.0020)	(0.0015)	(0.0810)	(0.0810)

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year dummies and number of high school graduates. Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

high school graduate regional level is our instrument for higher education attainment. From now we will refer this variable as *intensity of policy* or *available slots per high school graduate* in short. Also, note that whenever we use available slots per high school graduate to explain higher education attainment, we also control for number of high school graduates to exclude the effect of artificial increase in the instrument, resulted from demand side variations.

We also check the validity of our instrument for both males and females for whole and wage samples. The influence of available slots per high school graduate on higher education attainment is given in Table 4, with statistics for exclusion restriction test for both probit and OLS estimations. According to Angrist and Pischke (2009), our instrument has a statistically significant explanatory power for all sample specifications and passes instrument weakness test.

TABLE 4

	Dependent var: Higher Education Attainment			
	Whole Sample		Wage Sample	
	Males	Females	Males	Females
Intensity of Policy	0.077***	0.158***	0.066**	0.142***
	(0.025)	(0.027)	(0.031)	(0.044)
Observations	76,332	78,368	43,971	18,050
LR test, χ^2	9.92	33.73	4.37	10.65
LR test, P-value	(0.0016)	(0.0000)	(0.0365)	(0.0011)
F-test, by OLS	9.71	36.11	4.37	10.49
F-test, P-value	(0.0021)	(0.0000)	(0.0376)	(0.0014)

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year dummies and number of high school graduates. Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

5. Econometric Methodology

In this section, we explain the methods we use to identify the causal effect of education on wages and give our estimation results.

In order to identify wage returns to higher education, the joint model of education and wage equations is estimated by using available slots per high school graduates as an instrument for higher education. In these estimations, we use wage sample, which includes full-time workers (30-72 hours per week) older than 22 with observable wages. The simultaneous equations model is as follows:

$$h_{irt} = I(\gamma s_{irt} + X_{h,i}{}'\beta_h + \epsilon_{irt} \geq 0) \tag{5.1}$$

$$w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt} \tag{5.2}$$

where w_{irt} denotes the log of hourly wages, h_{irt} denotes binary higher education attainment variable, s_{irt} is the intensity of the exposure to policy at the regional level, the control variables $X_{h,i}{}'$ and $X_{w,i}{}'$ contains age and age square variables, region and cohort effects and year of survey dummies, and average income at the regional level (only for wage equation).

6. Returns to Education:Results

We first estimate two stage model (Equations 5.1 and 5.2) by using MLE estimation and the results for both males and females. We then estimate this two stage model including GDP per capita in a region to control for time varying region effects. We include this variable in order to address possible concerns that opening a university may affect labor market outcomes by improving regional economic opportunities in addition to improving the educational opportunities of individuals. Due to lack of GDP per capita data, we restrict data to cohorts 2004-2012 for these estimations. Next, we

control the robustness of our results by restricting our sample to the ones who did not migrate since their age 16, i.e. for education reasons, and estimate returns to education for those individuals, also with and without control for regional characteristics by GDP per capita.

Table 5 our first results from estimation of the two stage model (Equations 5.1 and 5.2) for 2002-2012 cohorts.

TABLE 5

	Male		**Female**	
	IV-MLE	OLS	IV-MLE	OLS
Higher Education Attainment (LATE)	0.483*** (0.012)	0.464*** (0.010)	0.136*** (0.058)	0.490** (0.014)
$\rho_{(\epsilon,\nu)}$	-0.028*** (0.007)		0.464*** (0.057)	
σ_ϵ	0.395*** (0.003)		0.472*** (0 .014)	
Observations	43,971	18,050	43,971	18,050

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year, average income and level of education dummies (only wage equation), intensity of policy and number of high school graduates (only education equation) . Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

The estimated returns to higher education is slightly higher than obtained by OLS estimation for males, while it is lower than OLS estimate for females. Males who attain higher education earn % 48 more, and females with higher education degree earn % 14 more compared to high school graduates. The returns to education for females are significantly lower than returns for males. Note that, IV estimations capture local average treatment effect for compliers, returns to education for those who attain education as a result of changes in the instrument. Thus, we can conclude that both males and females who attain higher education as a result of increasing access to universities also gain in terms of their wages.

The correlation between the error terms of education and wage equation is negative significant for males and positive significant for females. In other words, the unobserved characteristics which leads males to earn more as a college graduate are positively correlated with unobserved characteristics which leads them less likely to attain higher education. Thus, according to Table 5 while we observe negative selection of male compliers into education, we have evidence for positive selection of female compliers into education. The patterns of selection into education can be explained by the socio-economic characteristics of the country. Males are seen as bread-winners and encouraged to enter labor market as soon as possible. Thus, the ones with comparative advantage in labor market prefer to not attain education, and participate in labor force

right after high school graduation. This leads to negative selection of males into higher education. Conversely, parents tend to not allow their daughters to go to school, thus to be able to attain higher education females need to prove themselves to their teachers and their parents, which leads to positive selection of females into higher education (Caner et al., 2016).

An individual's decision of having higher education may also be affected by the characteristics and economic condition of the region he lives. One may argue that higher education expansion not only increases an individuals education opportunities but also her economic opportunities via positive effects on the local economy. Thus, we estimate the equation 4.1 by including region of residence's GDP (gross domestic product) per capita to in order to address the problem with unobserved heterogeneity due to omitted variable bias. Due to limitations of GDP per capita data at the region level, we restrict the wage sample to cohorts of 2004-2012 and estimate returns to education with and without GDP per capita control in the education equation to examine whether differences in coefficients are due to including GDP per capita or reduction in the number of cohorts.

TABLE 6

	Male		**Female**	
	(1)	(2)	(3)	(4)
Higher Education Attainment	0.541***	0.552***	0.054	0.062
	(0.017)	(0.017)	(0.099)	(0.102)
$\rho_{(\epsilon,\nu)}$	-0.042***	-0.058***	0.594***	0.587***
	(0.011)	(0.011)	(0.055)	(0.020)
σ_ϵ	0.399***	0.399***	0.507***	0.505***
	(0.004)	(0.004)	(0.055)	(0.055)
GDP per capita	No	Yes	No	Yes
Observations	31,742	31,742	13,449	13,449

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year, average income and level of education dummies (only wage equation), intensity of policy and number of high school graduates (only education equation) . Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

Table 6 reports our results for simultaneous equations model for cohorts 2004-2012 with and without GDP per capita control in the first stage. Our estimations suggest that, the magnitude of returns to higher education for females, decreases and its significance vanishes, when we restrict our sample to more recent cohorts exposed to expansion period, even before controlling for macroeconomic conditions in the first stage. For males, returns to higher education is significant and even higher when we restrict our sample to younger cohorts. For both males and females controlling for economic conditions increases the estimated returns to education in magnitude, and

does not change the patterns of significance. These results are consistent with the fact that mainly less developed regions are targeted during this expansion period, and so regions with low GDP per capita and low college graduates earlier to expansion, are the ones most of the compliers live. Thus, we estimate higher returns to education after controlling for regional economic characteristics.

In order to check the robustness of our results, we estimate the effect of policy on a subsample, called local sample, by restricting data to individuals who did not migrate after their age 16 until the survey year. This is a subset of sample of individuals who did not migrate to benefit the educational opportunities outside of their region of residence. Since the expansion has started with the establishment of new universities in cities which do not have one in it, these individual are the ones expected to be affected most from the expansion.

TABLE 7

	Male		**Female**	
	Higher Education	Log wages	Higher Education	Log wages
Intensity of Policy	0.098***		0.320***	
	(0.031)		(0.061)	
Higher Education Attainment		0.396***		0.835***
		(0.016)		(0.116)
$\rho_{(\epsilon,\nu)}$		-0.024***		-0.069***
		(0.025)		(0.085)
σ_ϵ		0.361***		0.403***
		(0.001)		(0.003)
Observations	31,384	31,384	11,180	11,180

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year, average income and level of education dummies (only wage equation), intensity of policy and number of high school graduates (only education equation) . Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

The results for both the impact of intensity of policy on higher education attainment and returns to education are reported in Table 7. Similarly, to the wage sample estimation results, intensity of policy has a significant and positive impact on educational attainment, with a higher effect for females. The returns to higher education for males in local sample are significant and positive but less than the estimated returns for males in wage sample. However, the estimated returns to education for females in local sample is positive and significant and also significantly higher than those for females in wage sample. This result coincides with the socio-economic characteristics of Turkey. Parents tend to not send their daughters to other cities for education purposes. The establishment of new universities all around the country, give females, who cannot take their parents' permission to attend higher education in other cities, are

now able to go to local universities. These individuals also still live in a city where the number of college graduates is low, due to non-existence of a university until recent expansion policy. Thus their returns to higher education is expected to be higher than wage sample results.

We also estimate negative correlation coefficient for both males and females, which suggests the negative selection of males and females among those who did not migrate since their age 16, into education. Thus, the unobserved characteristics which is correlated with high earnings with college degree is negatively correlated with unobserved characteristics which make individuals less likely to attain higher education for both genders. We expect to observe this negative ability bias among those who did not migrate for educational reasons, because the rapid establishment of new local universities signals some quality issues of these universities. Also we are looking at stayers here. Those with better job market opportunities elsewhere may have moved away.

TABLE 8

	Male		**Female**	
	(1)	(2)	(3)	(4)
Higher Education Attainment	0.395***	0.418***	0.585***	0.585***
	(0.020)	(0.025)	(0.064)	(0.068)
$\rho_{(\epsilon,\nu)}$	-0.061***	-0.099***	-0.306***	-0.309***
	(0.018)	(0.029)	(0.083)	(0.088)
σ_ϵ	0.357***	0.358***	0.410***	0.410***
	(0.003)	(0.003)	(0.011)	(0.011)
GDP per capita	No	Yes	No	Yes
Observations	23,249	23,249	8,545	8,545

Notes: Robust standard errors clustered at region·cohort level, are given in parentheses. Control variables: constant, age, age square variables, region fixed effects, cohort fixed effects, survey year, average income and level of education dummies (only wage equation), intensity of policy and number of high school graduates (only education equation) . Cohort 2008 is excluded. *p<0.1 **p<0.05 ***p<0.01.

We further analyze, returns to education for local sample by restricting the sample to cohorts 2004-2012, to be able to control for regional macroeconomic conditions. Similarly, we estimate joint model of education and wages (Equations 5.1 and 5.2) and the results are reported in Table 8. Even though, we show that there is no significant returns to higher education for females of younger cohorts with and without GDP per capita control for the wage sample, for the local sample females earn significantly higher than high school graduates. The returns to higher education is lower than wage sample but still positive and significant in the local sample even after controlling for time varying regional characteristics, and similar to our previous results returns are estimated to be higher after including GDP per capita to control variables of education equation.

Our results show that the estimated returns to education varies across different specifications and samples. In other words, the returns to education varies among compliers too. Hence, in future work, we plan to control for heterogeneous returns to education across individuals, in addition to different subsets of compliers.

7. Conclusion

In Turkey, higher education has expanded by the establishment of new universities and the increase in the available slots during 2006-2009 period. This expansion has been initiated by the central government with economic and political reasons, which creates an exogenous variation in the supply of higher education. We estimate the causal effect of higher education attainment on labor market outcomes, namely wages, by using this exogenous variation in educational opportunities at the regional level as an instrument for higher education attainment. This instrument is valid for various samples of young cohorts of age 23-32. The higher education attainment, for those who benefitted from the increase in available slots per high school graduates, leads to %48 and % 14 higher wages for men and women respectively, compare to high school graduates. We also show that the returns to higher education and selection of individuals into education varies a lot according to sample selection. This suggests the heterogeneity in returns to higher education and we aim to analyze this as a future work.

References

[1] Arap, K. (2010), Türkiye yeni üniversitelerine kavuşurken: Türkiye'de yeni üniversiteler ve kuruluş gerekçeleri, *Ankara Üniversitesi SBF Dergisi*, **65**(3), 001-029.

[2] Berlinski, S., Galiani, S. (2009), The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment, *Labour Economics*, **14**(3), 665-680.

[3] Blanden, J., Machin, S. (2004),Educational inequality and the expansion of UK higher education , *Scottish Journal of Political Economy*, **51**(2), 230-249.

[4] Brand, J. E., Xie, Y. (2010), Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education, *American sociological review*, **75**(2), 273-302.

[5] Bratti, M., Checchi, D., De Blasio, G. (2008), Does the expansion of higher education increase the equality of educational opportunities? Evidence from Italy, *Labour*, **22**(s1), 53-88.

[6] Caner A., Ökten Ç. (2010), Risk and Career Choice: Evidence from Turkey, *Economics of Education Review*, **29**(6), 1060-1075.

[7] Caner A., Ökten Ç. (2013), Higher education in Turkey: Subsidizing the rich or the poor?, *Economics of Education -Review*, **35**, 75-92.

[8] Caner A., Guven, C., Okten, C., Sakalli, S. O. (2016), Gender roles and the education gender gap in Turkey, *Social Indicators Research*, **129(3)**, 1231-1254.

[9] Card, D. (1995), Using geographic variation in college proximity to estimate the return to schooling , *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkammp, by LN.Christofides, E.K.Grant, R. Swidinsky.*, 201-222.

[10] Carneiro, P., Heckman, J. J., Vytlacil, E. J. (2011), Estimating marginal returns to education , *The American economic review*, **101**(6), 2754-2781.

[11] Deininger, K. (2003), Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda, *Economics of Education review*, **22**(3), 291-305.

[12] Dinçer, M. A., Kaushal, N., Grossman, M. (2014), Womens education: Harbinger of another spring? Evidence from a natural experiment in Turkey, *World Development*, **64**(3), 243-258.

[13] Di Paolo, Antonio and Tansel, Aysit (2017), Analyzing Wage Differentials by Fields of Study: Evidence from Turkey, *ERC Working Papers in Economics*, **14**(3), 17/08, July.
[14] Duflo, E. (2011), Consequences of school construction in Indonesia, *American Economic Review*, **91**(4), 795-813.
[15] Gurgand, M., Maurin, E. (2071), A large scale experiment: wages and educational expansion in France, *PSE Working Papers*.
[16] Kirdar, M. G., Dayoğlu, M., Koç, I. (2015), Does longer compulsory education equalize schooling by gender and rural/urban residence?, *The World Bank Economic Review*, **30**(3), 549-579.
[17] Kwiek, M. (2013), From system expansion to system contraction: Access to higher education in Poland, *Comparative Education Review*, **57**(3), 553-576.
[18] Kyui N. (2016), Expansion of higher education, employment and wages: Evidence from the Russian Transition, *Labour Economics*, **39**, 68-87.
[19] Maurin, E., McNally, S. (2008), Vive la revolution! Long-term educational returns of 1968 to the angry students, *Journal of Labor Economics*, **26**(1), 1-33.
[20] Oppedisano, V. (2011), The (adverse) effects of expanding higher education: Evidence from Italy, *Economics of Education Review*, **30**(5), 997-1008.
[21] Polat, S (2017), The expansion of higher education in Turkey: Access, equality and regional returns to education, *Structural Change and Economic Dynamics*, **43**, 1-14.
[22] Tansel A. (2010), Changing returns to education for men and women in a developing country: Turkey, 1994-2005, *In ESPE 2008 conference*.
[23] Tansel A., Bircan F. (2006), Demand for education in Turkey: A tobit analysis of private tutoring expenditures, *Economics of Education Review*, **25**(3), 303-313.
[24] Tsai, S. L., Xie, Y. ((2008), Heterogeneity in returns to college education: Selection bias in contemporary Taiwan, *Social Science Research*, **40**(3), 796-810.
[25] Walker, I., Zhu, Y. (2011), The college wage premium and the expansion of higher education in the UK, *The Scandinavian Journal of Economics*, **30**(5), 997-1008.
[26] Wang, X., Fleisher, B. M., Li, H., Li, S (2011) Access to college and heterogeneous returns to education in China, *Economics of Education Review*, **42**, 78-92.

8. Appendix

8.1. **Simultaneous equations model of educational attainment and wages.** In a traditional empirical study, returns to education are obtained by regressing wages w_{irt} for individual i, who lives in region r and from cohort t, on higher education attainment h_{irt} as follows:

$$w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}$$

However, these estimates are known to potentially suffer from ability bias, because of the correlation between unobserved characteristics ν_{irt} and higher education attainment h_{irt}. Thus, we define a simultaneous equations model of education and wages, by instrumenting higher education by exogeneous variations in education opportunities as follows:

$$h_{irt} = I(\gamma s_{irt} + X_{h,i}{}'\beta_h + \epsilon_{irt} \geq 0) \tag{8.1}$$

$$w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt} \tag{8.2}$$

where

$$\left\{ \begin{array}{c} \epsilon_{irt} \\ \nu_{irt} \end{array} \right\} \rightarrow N \left\{ E = \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \Sigma = \begin{pmatrix} {\sigma_\epsilon}^2 & \rho_{(\epsilon,\nu)} \\ \rho_{(\epsilon,\nu)} & 1 \end{pmatrix} \right\}$$

w_{irt} denotes the log of hourly wages,h_{irt} denotes binary higher education attainment, s_{irt} is the intensity of the exposure to policy at the regional level, $X_{h,i}{}'$ and $X_{w,i}{}'$ denoes control variables for education and wage equation respectively.

This model is the case of two equations with correlated random terms: a probit equation for educational attainment and a linear equation for wages. We estimate the parameters of the model $\theta = \{\beta_h, \beta_w, \gamma, \psi\}$ where the control variables of joint model are $X = \{X_{h,i}, X_{w,i}, h_{irt}\}$ by maximum likelihood estimation by maximizing the following likelihood function:

$$l(\theta, X_i) = \prod_{i=1}^{N} \{(l_{h_{irt}=1}(\theta, X))^{h_{irt}} (l_{h_{irt}=0}(\theta, X))^{1-h_{irt}}\}$$

where $l_{h_{irt}=1}(\theta, X)$ represents the likelihood function for individuals with higher education attainment and $l_{h_{irt}=0}(\theta, X)$ represents the likelihood function for individuals without higher education attainment. The details of likelihood functions are given as follows:

$$\begin{aligned}
l_{h_{irt}=1}(\theta, X) &= Pr(\gamma s_{irt} + X_{h,i}{}'\beta_h + \epsilon_{irt} \geq 0; w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}) \\
&= f(w_{irt}) Pr(\gamma s_{irt} + X_{h,i}{}'\beta_h + \epsilon_{irt} \geq 0 | w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}) \\
&= \frac{1}{\sigma_\nu} \phi(\frac{w_{irt} - \psi h_{irt} - X_{w,i}{}'\beta_w}{\nu})(1 - Pr(\epsilon_{irt} < -\gamma s_{irt} - X_{h,i}{}'\beta_h | w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}) \\
&= \frac{1}{\sigma_\nu} \phi(\frac{w_{irt} - \psi h_{irt} - X_{w,i}{}'\beta_w}{\nu})(1 - \Phi(\frac{-\gamma s_{irt} - X_{h,i}{}'\beta_h - \frac{\rho_{\epsilon\nu}}{{\sigma_\nu}^2}(w_{irt} - \psi h_{irt} - X_{w,i}{}'\beta_w)}{\sqrt{1 - \frac{\rho_{\epsilon\nu}{}^2}{{\sigma_\nu}^2}}}))
\end{aligned}$$

and

$$
\begin{aligned}
l_{h_{irt}=0}(\theta, X) &= Pr(\gamma s_{irt} + X_{h,i}{}'\beta_h + \epsilon_{irt} < 0; w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}) \\
&= f(w_{irt}) Pr(\gamma s_{irt} - X_{h,i}{}'\beta_h + \epsilon_{irt} < 0 | w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt}) \\
&= \frac{1}{\sigma_\nu}\phi(\frac{w_{irt} - \psi h_{irt} + X_{w,i}{}'\beta_w}{\nu})(Pr(\epsilon_{irt} < -\gamma s_{irt} - X_{h,i}{}'\beta_h | w_{irt} = \psi h_{irt} + X_{w,i}{}'\beta_w + \nu_{irt})) \\
&= \frac{1}{\sigma_\nu}\phi(\frac{w_{irt} - \psi h_{irt} - X_{w,i}{}'\beta_w}{\nu})(\Phi(\frac{-\gamma s_{irt} - X_{h,i}{}'\beta_h - \frac{\rho_{\epsilon\nu}}{{\sigma_\nu}^2}(w_{irt} - \psi h_{irt} - X_{w,i}{}'\beta_w)}{\sqrt{1 - \frac{{\rho_{\epsilon\nu}}^2}{{\sigma_\nu}^2}}}))
\end{aligned}
$$

where $\phi()$ denotes the density function and $\Phi()$ denotes the cumulative distribution function of standard normal distribution.