Education and American's Attitudes Towards Immigration

Juliet Yue

EC 15 Final Paper
December 14 2018

Abstract

This paper examines the effect of education on a person's attitudes towards immigration in the United States using survey data collected in the year of 2000. Two stage least squares (2SLS) methodology is applied, and father's highest degree is selected as the instrumental variable. The result indicated that for the population in the United States, there is a strong correlation between having a more open attitudes towards immigrants and participant's education attainment. The higher the degree individual has, the more likely for the individual to have a positive attitude towards immigrants. Secondly, the higher the degree is, the more positive impact it has on the individual's attitude.

1 Introduction

In the past decade, there has been a wave of anti-immigration sentiment growing in the US and across Europe (Grigorieff, Roth, and Ubfal 2018). We have witnessed the rise of far-right groups in Europe which manipulate and use the sentiment for their own benefits. Ever since Trump's presidential campaign, the wave has been hitting hard on Americans, affecting not only life of millions of immigrants, but also widen the divide between people in the United States. Thus, it is important for us to understand the factors play behind such a sentiment and identify key characteristics of the cause in order to device better public policies to ease the growing divide in America and Europe.

Education has long been identified as one of the key factors correlated with individual's attitudes towards immigration. Government agencies like the British Department of Business & Skills has conducted research studying the effect of higher education on graduates' attitudes over six major aspects, among which is graduates' attitudes towards immigration in the UK (Brennan et al. 2015). They found strong evidence supporting that people who have received a higher degree have a warmer attitudes towards immigration. It will be of our interest to see if such a trend is also applicable to the United States. To study the question, I obtain data from General Social Survey conducted in the United States, and choose Two stage least squares as my research methodology as there are some potential issues with conducting only ordinary least square regression. For example, a person's education attainment may be affected by his/her parent's education attainment. A person may obtain a warmer attitude towards immigration if his/her father have a higher degree because father's higher degree suggests that the father may have take a more welcoming stand on immigration and pass that attitude down on his children. Omitting father's degree level may result in an upward bias on the coefficient on education attainment. One thing to note about this research is that the data used in this study comes from only the year 2000. So the exact magnitude of the coefficients may not be representative anymore; however the general trend is expected to have remained the same.

2 Literature Review

There have been many published economics research papers on the effect of education on individual's attitude towards immigration and all three papers discussed in this section studied (at least partially) the same question as this paper. In 2012, Melgar conducted a study to understand the impact of education on attitudes towards immigrants and free-Trade. The research methodology used is Difference in Difference with interaction terms between education level and respondent's country's GDP per capita. Melgar hypothesized that high educated people in richer countries are more likely to support immigrants. While the result confirmed that the educational level has a positive impact on attitude towards immigration, it also indicated that high educated people in richer countries could also be against immigration if their overall attitude to these issues is bad (if their nationalism propensity is high) (Melgar 2012). Melgar included variable indicating a person's attitude on nationalism, which is not an observable factor in my study. Thus, there may be omitted variable bias in the result I present because of this.

Hainmueller and Hiscox conducted study in Europe with the main interest to see whether "labor-market competition was a potent source of anti-immigrant sentiment, in particular among less-educated or less-skilled citizens who fear being forced to compete for jobs with low-skilled immigrants willing to work for much lower wages" (Hainmueller and Hiscox 2007). Result showed that people with higher levels of education were more likely to favor immigration regardless of the skill attributes of the immigrants. The average occupational skill set or education level of immigrants in a country did not have a significant impact on non-immigrant's attitudes on the matter, and whether individual was competing for work in the labor market has no significant impact on individual's attitude towards immigration holding all else equal.

The third paper by Scheve and Slaughter used data from National Election Studies surveys conducted in the United States. They studied the impact on individual's preferences over immigration policy by regressing attitudes on years of education and occupation wage respectively using data from 1992, 1994, and 1996. They also used

difference in difference to test the impact of respondent's geographical region interacting with both independent variables of interest. Result suggested that "workers with less years of education are significantly more likely to prefer limiting immigrant inflows into the United States" and there was no evidence suggesting that the relationship between years of education and immigration opinions is stronger in high immigration communities (Scheve and Slaughter 2001).

3 Data Description and Summary Statistics

Data used in this research come exclusively from General Survey Data conducted in the year 2000. Observations with missing critical data (education levels) are dropped from the data set so that missing data does not interfere with the result. I create a variable to represent a person's attitude towards immigration using the LETIN variable from GSS. The question asks the respondent's opinion on if they think more foreign citizens should be permitted to work in the United states. LETIN takes the value of 1-5 where 1 being the number should increase a lot and 5 being the number should decrease a lot. Before conducting further tests on the data set, I first conduct a simple balance test to compare respondent who has answered the question and respondent who has a missing value on the question due the worry that there may be unobserved selection bias on people who choose to answer the question and those who choose not to answer the question. I create a variable to indicate female respondents, a variable to represent if respondent indicates that he/she is working full-time/part-time. The result of the preliminary balance test is presented in Table 1 and from the t-test, no coefficient is statistically significant so I can be certain that there is not selection bias on this set of data with people with non-missing value for LETIN. All observations with missing value are then dropped, which leaves us in total 972 observations as our testing data set.

Reviewing the history of immigration policy in the United States, I find that around the year 2000, the government had a relatively open attitudes towards immigration. Thus, I define the dummy variable welcome as 1 when respondent chooses options 1-3 where 3 is remain the same level as now; welcome takes a value 0 when respondent chooses options

4 and 5.

I also create a few more variables: a variable to indicate if a respondent has a bachelor or higher degree, a variable to indicate if a respondent has a graduate or higher degree, and categorize regions where respondent was at at the age 16 into 4 major areas in the US and foreign country. Descriptive statistics is shown in Table 2. From the table, there is no significant increase for the means of years of education between the two groups. While on means of degree levels, including respondent's degree, respondent's father's degree, respondent who has a bachelor degree or higher and respondent who has a graduate degree or higher, I observe that there is a more significant increase. For example, on the variable bachelor degree or higher, the mean for respondent who is positive about immigrants is twice in number than those who is negative about immigrants. On the variable graduate degree or higher, the mean for respondent who is positive about immigrants is 4 times larger than that of the group of respondents who do not welcome immigrants. One surprising thing is that the total family income level as well as respondent's work status seem to have no difference between the two groups, both of which I assumed may affect individual's attitude towards immigration.

4 Model

In order to understand the effect of education on attitude towards immigration I choose the two stage least square methodology because ordinary least square test may be affected by omitted variable bias as discussed in the introduction section. The dependent variable(Y) I want to study is the attitude towards immigration. Because there are many ways to manipulate data on education, I select 4 different critical independent variables all representing respondent's education level (years of education, degree level, if respondent has a bachelor degree or more, if a respondent has a graduate degree or more) to have a comprehensive understanding of the matter. I choose respondent's father's degree level as instrumental in all tests. The general OLS (naive) regression is represented as below:

$$\text{welcome} = \beta_0^{naive} + \beta_1^{naive}\text{Education level} + \delta \text{R}egion_i + \tau \text{X} + \mu \tag{1}$$

I include geographic region where respondent grow up in to account for region fixed effect. I also include age, female, work status, and total family income level as my covariates represented by X in the equation. These are the variables that I think will control the permanent difference among individuals. For example, respondent who grows up in a region that has larger immigrant population may have a more open attitude towards immigration. Older people may have a more negative attitude towards immigration.

For 2SLS tests, the general equations for first stage regression, reduced form, and 2SLS regression are shown below:

Reduced form: (same across all four tests)

$$\text{welcome} = \beta_0^{RF} + \beta_1^{RF}\text{Father's Degree} + \Phi \text{R}egion_i + \omega \text{X} + \epsilon \tag{2}$$

First Stage:

$$\text{Education level} = \Pi_0 + \Pi_1 \text{Father's Degree} + \gamma \text{R}egion_i + \alpha \text{X} + \varepsilon \tag{3}$$

2SLS:

$$\text{welcome} = \beta_0^{2SLS} + \beta_1^{2SLS}\text{Education level} + \delta' \text{R}egion_i + \tau' \text{X} + \mu' \tag{4}$$

Dependent variable welcome takes a value of 1 if respondent report positive attitude on immigration and 0 if negative. Positive Coefficient estimate suggests that the variables have a positive impact on person's attitude on immigration and vice versa.

The first test takes respondent's years of education to represent individual's education level. The second test uses respondent's highest degree as the endogenous variable. The advantage of using degree over years of education is that one additional year of education may has less impact on individual's overall life comparatively. For example, we often observe a huge jump on college student's income before and after the student graduates whereas it is very common for a college freshman to earn about the same as a college sophomore or junior.

Another way to represent individual's education level is to use dummy variables. I separate the group in to those who have obtained higher education (=1 if bachelor degree or above/graduate degree or above) and those who have not correspondingly with the

dummy variables created. The third test takes bachelor degree or more as endogenous variable, and the fourth test takes graduate or more as endogenous variable. The results can give us a closer look into the effect of higher education on respondent's attitude towards immigration.

For all of the tests above, I expect to see a positive coefficient estimates on respective endogenous variable. I also expect that the coefficient estimate on degree to be larger than that of years of education. One additional degree should have a larger impact than one additional year of education for the respondent. The coefficient estimate on graduate or more is expected to be larger than that of bachelor or more. The higher the respondent's education level is, the more likely that the respondent welcomes immigrants compare to the rest of the population.

5 Empirical Analysis

With the models explained above, the result on main variable of interest is presented in Table 3. All coefficients estimates in First Stage regressions are statistically significant at 1 percent level. This means that the instrumental variables are highly correlated with the endogenous variables, which satisfies the first assumption for instrumental variable. In terms of the independence assumption, some compromise may need to be made here as father's education level may be correlated with omitted factors like race. Lastly, of all the covariates I have included in the model, I assume that father's education level has the most direct impact on respondent's education level and have little or no impact to respondent's attitude towards immigration through other channels. The selection of instrumental variable is certainly not perfect, but at least given the result from first stage regression, we know that the instruments all have strong casual effects on the endogenous variables.

The 2SLS test on the effect of individual's years of education shows that the coefficient estimate is statistically significant at 5 percent level. One additional year of education increases individual's attitude towards immigration by 3.6 percentage point holding all else equal. It is about on par with the OLS regression (3.4 percentage point

increase).

In the second 2SLS regression on attitude towards immigration and individual's degree level, using father's degree level as instrument, the result indicates a statistically significant coefficient estimates on degree. One additional degree obtained increases individual's attitude towards immigration by 8.1 percentage point (significant at 5 percent level), which is roughly the same as the result (7.6 percentage point) indicated by OLS (significant at 1 percent level), holding all else equal.

The third row shows the regression on whether respondent has a degree of bachelor or more with father's degree as instrument. On average, from 2SLS, one additional degree warms up respondent's attitude towards immigration by 23.9 percentage point (significant at 5 percent level), up from 18.7 percentage point indicated by OLS regression (significant at 1 percent level), holding all else equal.

The fourth row shows the regression on whether respondent has a degree of graduate or more with father's degree as instrument. One additional degree increases the likelyhood for respondent to welcome immigration by 49.3 percentage point (significant at 5 percent level), which is up from 30.0 percentage point indicated by OLS regression (significant at 1 percent level), holding all else equal.

Table 4 shows the coefficient estimates on all covariates in each stage of the 2SLS test when degree is the endogenous variable. The omitted variable here is foreign, indicating that the respondent resided in a foreign country at the age of 16, a group which I expect to have the highest acceptance of immigration and the result indeed proves so (constant term has a value as high as 0.938). I use this table to illustrate the general effect of the covariates. Compare to respondents who resided in a foreign country at the age of 16, respondents who resided in all regions in the US has a less welcoming attitude towards immigration (all coefficients significant at 1 percent level except for northeast). Family total income level (significant at 10 percent level) has a slight negative effect on attitude towards immigration which is a bit counter intuitive, and I am interested in learning more about the possible explanations to it. Age, gender and work status has no

significant effect on individual's attitude towards immigration.

The result is what I have expected – education level is positively correlated with individual's attitude towards immigration. Moreover, given the disparities of the 2SLS coefficient estimates from the three regressions based on respondent's degree level, it also shows that the effect of degree level is not evenly distributed. The higher the degree is, the more impact it has on a attitude towards immigration. One possible explanation is that higher education level makes people more informed of the benefits immigrants bring to the country, the actual impact of immigration policy to the country, and have a generally more open attitudes towards other cultures. With all these factors combined, the higher the person's education level is, the more open attitudes this person will hold towards immigration.

6 Conclusion

From the result of my statistical analysis, it is safe to conclude that the higher the person's education level is, the more open attitudes this person will hold towards immigration. Such a result coincides with the findings from all three papers mentioned in the literature review section. I can also conclude that the higher the degree level is, the more impact one additional degree has on individual's attitude towards immigration. Additionally, like Hainmueller and Hiscox, my result also shows that the work status of an individual has no significant impact on respondent's attitude towards immigration.

Scheve and Slaughter found that there was no evidence suggesting that the relationship between years of education and immigration opinions is stronger in high immigration communities (2001). In my research, I account for region fixed effect by controlling the region respondent was at at age of 16, and have observed statistically significant difference across different regions on the impact towards respondent's attitude. Such a contradictory finding is perhaps due to the difference methodologies in categorizing region. Scheve and Slaughter isolated specifically metropolitan statistical areas that have high immigrant populations as the geographic control whereas I only categorize geographical locations into 5 categories and there may be a lot of variations within each region and other per-

manent differences are absorbed by it. It is still valid to say that there is permanent difference on attitudes towards immigration across each region but I cannot know if such difference is caused by the percentage of immigrant population in each region.

The study can be improved by incorporating race factors into the regression as race may be an important factor that impact individual's attitude over immigrants. Additionally, individual's nationalism propensity is not included in the study, which, according to the study conducted by Melgar, can also result in omitted variable bias. More recent data, and data from multiple years can be included to further our understanding on the topic.

Moving forward, from the result of the study as well as previous research on the subject matter, I believe the the reason why we can repeatedly observe a strong correlation between higher education and attitude towards immigration is because education makes people more informed of the benefits immigrants bring to the country, the actual impact of immigration policy to the country, and have a generally more open attitudes towards other cultures. It is hard to solve the problem of education resource disparity in one day, but we can start putting more effort in informing the overall population on the actual impact of immigration on the country as a whole. We should also advocate for cultural diversity in our society, encourage communications and interactions between immigrant population and non-immigrant population in order to reduce xenophobia in the country.

References

Brennan, John, Jenny Chanfreau, Jerome Finnegan, Julia Griggs, Zsolt Kiss, and Alison Park. 2015. "The effect of higher education on graduates' attitudes: Secondary Analysis of the British Social Attitudes Survey." *BIS RESEARCH PAPER NO. 200.*

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal. 2018. "Does Information Change Attitudes Towards Immigrants? Representative Evidence from Survey Experiments." Working paper, March.

Hainmueller, Jens, and Michael J. Hiscox. 2007. "Educated Preferences: Explaining Attitudes Toward Immigration in Europe." *International Organization* 61 (2): 399–442.

Melgar, Natalia. 2012. "Interaction Effects in Probit Models, Reinterpreting the Impact of Education on Attitudes Towards Immigrants and Free-Trade." *Journal of Reviews on Global Economics* 1:82–88.

Scheve, Kenneth F., and Matthew J. Slaughter. 2001. "Educated Preferences: Explaining Attitudes Toward Immigration in Europe." *The Review of Economics and Statistics* 83 (1): 133–145.

Tables

Table 1: Balance test on Respondents who answered the question LETIN

	Respondent who answer the question on LETIN (1)	Respondent who does not answer the question on LETIN (2)	t-test
R's highest degree level	1.47	1.45	0.02
	(1.14)	(1.17)	(0.04)
	1310	1489	2799
R's father's highest degree level	1.01	1.04	-0.03
	(1.18)	(1.23)	(0.05)
	978	1087	2065
Group in which R's total family income level belongs	10.73	10.69	0.04
	(2.41)	(2.33)	(0.10)
	1172	1284	2456
age	45.41	46.56	-1.15
	(17.35)	(17.37)	(0.66)
	1313	1496	2809
female	0.56	0.57	-0.01
	(0.50)	(0.50)	(0.02)
	1318	1499	2817
Region where R was at at age 16	4.45	4.29	0.15
	(2.63)	(2.60)	(0.10)
	1318	1499	2817
Work Status (=1 if R is working full/part time)	0.35	0.36	-0.01
	(0.48)	(0.48)	(0.02)
	1318	1499	2817

Column 1 is the statistic summary for respondents who answered the question on LETIN, column 2 is the statistic summary for respondents who has missing value on the question on LETIN, column 3 contains the t-test value on each variable's coefficient when regress the variable on notmissing, which indicates whether LETIN is missing or not. In column 1 and 2, the first row indicates the mean, standard deviation is in parenthesis on the second row, and number of observations is on the third row. Asterisks are used to note statistical significance and no asterisk is present in this table.

Table 2: Descriptive Statistics

	Welcome immigrants	Do not welcome immigrants
Years of education	12.73	12.79
	(2.96)	(2.53)
	555	580
Highest Degree	1.63	1.27
	(1.24)	(0.96)
	555	575
Father's Highest Degree	1.11	0.88
	(1.25)	(1.06)
	555	423
Bachelor Degree or More	0.28	0.15
	(0.45)	(0.36)
	555	417
Graduate Degree or More	0.122	0.03
	(0.33)	(0.17)
	555	417
Total Family Income Level	10.71	10.77
	(2.42)	(2.39)
	498	372
Age	44.77	0.56
	(16.89)	(0.50)
	554	414
Female	0.55	0.56
	(0.50)	(0.50)
	555	417

Table 2 continued from previous page

	Welcome immigrants	Do not welcome immigrants
	0.34	0.35
Has work	(0.47)	(0.48)
	555	417
	0.22	0.18
Northeast	(0.41)	(0.38)
	555	417
	0.10	0.02
Foreign	(0.30)	(0.12)
	555	417
	0.25	0.26
Midwest	(0.44)	(0.44)
	555	417
	0.26	0.39
South	(0.44)	(0.36)
	555	417
	0.17	0.15
West	(0.37)	(0.36)
	555	417

Column 1 gives summary statistics on each variables for respondents who welcome immigrants. Column 2 also gives summary statistics on each variables but for respondents who do not welcome immigrants. In each cell, the first row indicates the mean, standard deviation is in parenthesis on the second row, and number of observations is on the third row.

Table 3: Effect of education on participants attitudes towards immigration

	Controls	All Participants			
		First stage (1)	Reduced form (2)	2SLS (3)	OLS . (4)
Welcome Immigrants (=1)	Years of education	0.885***	0.032**	0.036**	0.032***
		(0.075)	(0.015)	(0.016)	(0.005)
	All Degrees	0.376***	0.032**	0.081**	0.076***
		(0.031)	(0.015)	(0.038)	(0.013)
	Bachelor or more	0.133***	0.032**	0.239**	0.187***
		(0.012)	(0.015)	(0.109)	(0.034)
	Graduate or more	0.065***	0.032**	0.493**	0.300***
		(0.008)	(0.015)	(0.226)	(0.051)
Total observations		868			

Notes: (1)-(4) shows the result from First Stage, Reduced Form, 2SLS, and OLS regressions. In each cell, the first row indicates the coefficient estimates on the variable, standard deviation is in parenthesis on the second row. The number of observations in all regressions is shown on the bottom row.
*Significant at 10 percent; **significant at 5 percent; ***significant at 1 percent.

Table 4: Regression outcome using degree including all covariates

	First Stage	Reduced form	OLS	2SLS
degree			0.076***	0.081**
			(0.013)	(0,038)
father degree	0.376***	0.032**		
	(0.253)	(0.015)		
age	0.0063***	-0.0002	-0.001	-0.001
	(0.0024)	(0.0011)	(0.001)	(0.001)
total income level	0.090***	-0.009	-0.011*	-0.012*
	(0.017)	(0.077)	(0.006)	(0.009)
northeast	-0.038	-0.199*	-0.219***	-0.192
	(0.166)	(0.078)	(0.064)	(0.077)
midwest	-0.273*	-0.348***	-0.312***	-0.326***
	(0.159)	(0.074)	(0.062)	(0.074)
south	-0.319*	-0.475***	-0.416***	-0.448***
	(0.157)	(0.073)	(0.061)	(0.075)
west	-0.366**	-0.355***	-0.294***	-0.318***
	(0.167)	(0.078)	(0.065)	(0.078)
female	0.095	-0.012	-0.028	-0.020
	(0.716)	(0.033)	(0.029)	(0.033)
haswork	-0.299***	0.009	0.051	0.033
	(0.085)	(0.040)	(0.034)	(0.040)
constant	0.230	0.992***	0.938***	0.977***
	(0.253)	(0.118)	(0.096)	0.112
Observations	868	868	868	868

Notes: Column 1-4 shows the result from First Stage, Reduced Form, 2SLS, and OLS regressions for each variables. In each cell, the first row indicates the coefficient estimates on the variable, standard deviation is in parenthesis on the second row. Total number of observations in each regression is indicated on the last row.
*Significant at 10 percent; **significant at 5 percent; ***significant at 1 percent.

Stata Output and log file

```
* Juliet Yue
* EC 15 final project
* Prof. McInerney

* Set up the environment
clear all
capture log close

log using C:\Users\qyue01\Box\final.log, replace

use C:\Users\qyue01\Box\gss.dta

/*
 * variables that need to take into considerations
 * Y = attitude towards immigration (letin)
 * X = individual highest education achievement (degree/educ/bachmore/gradmore)
 * Z = father's education achievement (padeg)
 * Survey on Y is only conducted in the year 2000
 */

keep if year == 2000

* Respondents who has non-missing data on the varibale letin
gen notmissing = letin <= 5

*dummy variable for female
gen female = sex==2

* Identify respondents who has part-time/full-time work (=1)
gen haswork = hrs1 >89

* Balance test.
* Check if there is selection bias on those who has data on letin
summ degree padeg income age female reg16 haswork if notmissing == 1
summ degree padeg income age female reg16 haswork if notmissing == 0
reg degree notmissing
reg padeg notmissing
reg income notmissing
reg age notmissing
reg female notmissing
reg reg16 notmissing
```

```
reg haswork notmissing

* Drop observation with missing data on letin
drop if notmissing == 0
count if notmissing == 1

* Dummy variable for R's attitude towards immigration (=1 welcome)
gen welcome = letin <= 3

* Dummy variables for different regions in America
gen northeast = reg16==1 | reg16==2
gen midwest = reg16==3 | reg16 == 4
gen south = reg16==6 | reg16==7 | reg16==5
gen west = reg16==8 | reg16 == 9
gen foreign = reg16==0

* Drop observations with missing education data
drop if degree > 4
drop if padeg > 4

* Count all valid obervations
count if notmissing == 1

gen bachmore = degree==3 | degree==4
gen grad = degree==4

*Descriptive Stats
summ educ degree padeg bachmore grad age income northeast foreign midwest south west female haswork if welcome == 1
summ educ degree padeg bachmore grad age income northeast foreign midwest south west female haswork if welcome == 0

*Years of education
* Naive regression
reg welcome educ age income northeast midwest south west female haswork
* Reduced form
reg welcome padeg age income northeast midwest south west female haswork
* 2SLS
ivreg welcome (educ=padeg) age northeast midwest south west female haswork income, first

* Degree
* Naive regression
reg welcome degree age income northeast midwest south west female haswork
* Reduced form is the same as above
```

```
* 2SLS
ivreg welcome (degree=padeg) age northeast midwest south west female haswork income, first

* Bachelor or more
* Naive regression
reg welcome bachmore age income northeast midwest south west female haswork
* Reduced form is the same as above
* 2SLS
ivreg welcome (bachmore=padeg) age northeast midwest south west female haswork income, first

* Grad level or more
reg welcome grad age income northeast midwest south west female haswork
* Reduced form is the same as above
* 2SLS
ivreg welcome (grad=padeg) age northeast midwest south west female haswork income, first


log close



-------------------------------------------------------------------------------
      name:  <unnamed>
       log:  C:\Users\qyue01\Box\final.log
  log type:  text
 opened on:  14 Dec 2018, 16:58:21

.
. use C:\Users\qyue01\Box\gss.dta

.
. /*
>  * variables that need to take into considerations
>  * Y = attitude towards immigration (letin)
>  * X = individual highest education achievement (degree/educ/bachmore/gradmor
> e)
>  * Z = father's education achievement (padeg)
>  * Survey on Y is only conducted in the year 2000
>  */
.
. keep if year == 2000
(21,533 observations deleted)
```

```
.
. * Respondents who has non-missing data on the varibale letin
. gen notmissing = letin <= 5

.
. *dummy variable for female
. gen female = sex==2

.
. * Identify respondents who has part-time/full-time work (=1)
. gen haswork = hrs1 >89

.
. * Balance test.
. * Check if there is selection bias on those who has data on letin
. summ degree padeg income age female reg16 haswork if notmissing == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      degree |      1,310    1.471756    1.139758          0          4
       padeg |        978     1.01227    1.179774          0          4
      income |      1,172    10.73464    2.405245          1         12
         age |      1,313    45.40899    17.35008         18         89
      female |      1,318    .5569044    .4969399          0          1
-------------+---------------------------------------------------------
       reg16 |      1,318    4.446131     2.62655          0          9
     haswork |      1,318    .3467375     .476112          0          1

. summ degree padeg income age female reg16 haswork if notmissing == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      degree |      1,489    1.451981    1.173983          0          4
       padeg |      1,087    1.039558    1.227861          0          4
      income |      1,284     10.6947    2.332849          1         12
         age |      1,496    46.56083    17.36757         18         89
      female |      1,499    .5697131    .4952815          0          1
-------------+---------------------------------------------------------
       reg16 |      1,499    4.293529    2.601587          0          9
     haswork |      1,499    .3615744    .4806167          0          1
```

```
. reg degree notmissing

      Source |       SS           df       MS      Number of obs   =     2,799
-------------+----------------------------------   F(1, 2797)      =      0.20
       Model |  .272505595         1  .272505595   Prob > F        =    0.6522
    Residual |  3751.27162     2,797  1.34117684   R-squared       =    0.0001
-------------+----------------------------------   Adj R-squared   =   -0.0003
       Total |  3751.54412     2,798   1.3407949   Root MSE        =    1.1581

------------------------------------------------------------------------------
      degree |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |   .0197745   .0438694     0.45   0.652    -.0662451    .1057942
       _cons |   1.451981   .0300121    48.38   0.000     1.393133    1.510829
------------------------------------------------------------------------------

. reg padeg notmissing

      Source |       SS           df       MS      Number of obs   =     2,065
-------------+----------------------------------   F(1, 2063)      =      0.26
       Model |  .383360182         1  .383360182   Prob > F        =    0.6075
    Residual |  2997.15175     2,063  1.45281229   R-squared       =    0.0001
-------------+----------------------------------   Adj R-squared   =   -0.0004
       Total |  2997.53511     2,064  1.45229414   Root MSE        =    1.2053

------------------------------------------------------------------------------
       padeg |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |  -.0272885   .0531228    -0.51   0.608    -.1314683    .0768913
       _cons |   1.039558   .0365586    28.44   0.000     .9678627    1.111254
------------------------------------------------------------------------------

. reg income notmissing

      Source |       SS           df       MS      Number of obs   =     2,456
-------------+----------------------------------   F(1, 2454)      =      0.17
       Model |  .977300107         1  .977300107   Prob > F        =    0.6763
    Residual |  13756.7975     2,454  5.60586697   R-squared       =    0.0001
-------------+----------------------------------   Adj R-squared   =   -0.0003
       Total |  13757.7748     2,455   5.6039816   Root MSE        =    2.3677

------------------------------------------------------------------------------
```

```
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |   .0399376   .0956509     0.42   0.676    -.1476273    .2275025
       _cons |    10.6947   .0660753   161.86   0.000     10.56514    10.82427
------------------------------------------------------------------------------

. reg age notmissing

      Source |       SS           df       MS      Number of obs   =     2,809
-------------+----------------------------------   F(1, 2807)      =      3.08
       Model |  927.748517         1  927.748517   Prob > F        =    0.0794
    Residual |  845885.839     2,807  301.348713   R-squared       =    0.0011
-------------+----------------------------------   Adj R-squared   =    0.0007
       Total |  846813.587     2,808   301.57179   Root MSE        =    17.359

------------------------------------------------------------------------------
         age |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |  -1.151842   .6564662    -1.75   0.079    -2.439047    .1353633
       _cons |   46.56083   .4488166   103.74   0.000     45.68079    47.44087
------------------------------------------------------------------------------

. reg female notmissing

      Source |       SS           df       MS      Number of obs   =     2,817
-------------+----------------------------------   F(1, 2815)      =      0.47
       Model |  .115064967         1  .115064967   Prob > F        =    0.4941
    Residual |  692.697147     2,815  .246073587   R-squared       =    0.0002
-------------+----------------------------------   Adj R-squared   =   -0.0002
       Total |  692.812212     2,816  .246027064   Root MSE        =    .49606

------------------------------------------------------------------------------
      female |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |  -.0128087   .0187313    -0.68   0.494    -.0495372    .0239197
       _cons |   .5697131   .0128124    44.47   0.000     .5445904    .5948359
------------------------------------------------------------------------------

. reg reg16 notmissing

      Source |       SS           df       MS      Number of obs   =     2,817
-------------+----------------------------------   F(1, 2815)      =      2.39
```

```
       Model |   16.332313         1   16.332313   Prob > F        =    0.1221
    Residual |  19224.5225     2,815  6.82931527   R-squared       =    0.0008
-------------+----------------------------------   Adj R-squared   =    0.0005
       Total |  19240.8548     2,816  6.83268992   Root MSE        =    2.6133

------------------------------------------------------------------------------
       reg16 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |   .1526015   .0986787     1.55   0.122    -.0408884    .3460914
       _cons |   4.293529   .0674975    63.61   0.000     4.161179    4.425879
------------------------------------------------------------------------------

. reg haswork notmissing

      Source |       SS           df       MS      Number of obs   =     2,817
-------------+----------------------------------   F(1, 2815)      =      0.67
       Model |  .154389104         1  .154389104   Prob > F        =    0.4116
    Residual |  644.567656     2,815  .228976077   R-squared       =    0.0002
-------------+----------------------------------   Adj R-squared   =   -0.0001
       Total |  644.722045     2,816   .22894959   Root MSE        =    .47851

------------------------------------------------------------------------------
     haswork |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  notmissing |  -.0148369   .0180688    -0.82   0.412    -.0502664    .0205926
       _cons |   .3615744   .0123593    29.26   0.000     .3373402    .3858086
------------------------------------------------------------------------------

.
. * Drop observation with missing data on letin
. drop if notmissing == 0
(1,499 observations deleted)

. count if notmissing == 1
  1,318

.
. * Dummy variable for R's attitude towards immigration (=1 welcome)
. gen welcome = letin <= 3

.
. * Dummy variables for different regions in America
```

```
. gen northeast = reg16==1 | reg16==2

. gen midwest = reg16==3 | reg16 == 4

. gen south = reg16==6 | reg16==7 | reg16==5

. gen west = reg16==8 | reg16 == 9

. gen foreign = reg16==0

.
. * Drop observations with missing education data
. drop if degree > 4
(8 observations deleted)

. drop if padeg > 4
(338 observations deleted)

.
. * Count all valid obervations
. count if notmissing == 1
  972

.
. gen bachmore = degree==3 | degree==4

. gen grad = degree==4

.
. *Descriptive Stats
. summ educ degree padeg bachmore grad age income northeast foreign midwest sou
> th west female haswork if welcome == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        educ |        555    13.94414    2.953695          2         20
      degree |        555    1.720721    1.255825          0          4
       padeg |        555    1.111712    1.252488          0          4
    bachmore |        555    .3117117    .4636106          0          1
        grad |        555    .1351351    .3421763          0          1
-------------+---------------------------------------------------------
         age |        554    44.77256    16.89213         18         89
```

```
      income |        498    10.84337    2.353893          1         12
   northeast |        555    .2198198     .414498          0          1
     foreign |        555    .0972973    .2966295          0          1
     midwest |        555    .2648649    .4416593          0          1
-------------+--------------------------------------------------------
       south |        555    .2504505    .4336633          0          1
        west |        555    .1675676    .3738188          0          1
      female |        555    .5441441    .4984968          0          1
     haswork |        555    .3351351    .4724635          0          1

. summ educ degree padeg bachmore grad age income northeast foreign midwest sou
> th west female haswork if welcome == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        educ |        417    13.01679     2.51002          3         20
      degree |        417    1.354916    .9848277          0          4
       padeg |        417    .8872902    1.069353          0          4
    bachmore |        417    .1702638    .3763159          0          1
        grad |        417    .0383693    .1923169          0          1
-------------+--------------------------------------------------------
         age |        414    46.76329    17.68317         19         89
      income |        372    10.95968    2.159662          1         12
   northeast |        417    .1558753    .3631728          0          1
     foreign |        417    .0095923    .0975866          0          1
     midwest |        417    .2781775    .4486395          0          1
-------------+--------------------------------------------------------
       south |        417    .4028777    .4910657          0          1
        west |        417    .1534772    .3608799          0          1
      female |        417    .5443645    .4986261          0          1
     haswork |        417     .352518    .4783281          0          1

.
. *Years of education
. * Naive regression
. reg welcome educ age income northeast midwest south west female haswork

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     10.01
       Model |  20.1671607         9  2.24079564   Prob > F        =    0.0000
    Residual |  192.113945       858  .223909027   R-squared       =    0.0950
-------------+----------------------------------   Adj R-squared   =    0.0855
```

```
       Total |  212.281106       867  .244845566   Root MSE        =    .47319

------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0291219   .0061205     4.76   0.000     .0171089    .0411348
         age |  -.0004861   .0010487    -0.46   0.643    -.0025443    .0015722
      income |  -.0158029   .0077854    -2.03   0.043    -.0310835   -.0005222
   northeast |  -.1944783   .0766713    -2.54   0.011    -.3449636   -.0439931
     midwest |  -.3348486   .0732704    -4.57   0.000    -.4786587   -.1910384
       south |  -.4507028   .0727111    -6.20   0.000    -.5934154   -.3079903
        west |  -.3333631   .0770299    -4.33   0.000    -.4845521   -.1821741
      female |   -.019586   .0330225    -0.59   0.553    -.0844004    .0452284
     haswork |    .030565   .0393235     0.78   0.437    -.0466167    .1077466
       _cons |   .6982884   .1352286     5.16   0.000     .4328708    .9637061
------------------------------------------------------------------------------

. * Reduced form
. reg welcome padeg age income northeast midwest south west female haswork

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      7.82
       Model |  16.0865411         9  1.78739345   Prob > F        =    0.0000
    Residual |  196.194565       858  .228664994   R-squared       =    0.0758
-------------+----------------------------------   Adj R-squared   =    0.0661
       Total |  212.281106       867  .244845566   Root MSE        =    .47819

------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       padeg |   .0303511   .0145977     2.08   0.038     .0016998    .0590024
         age |  -.0003694   .0010989    -0.34   0.737    -.0025262    .0017875
      income |  -.0085293   .0076771    -1.11   0.267    -.0235974    .0065387
   northeast |   -.194673   .0774985    -2.51   0.012    -.3467818   -.0425642
     midwest |  -.3479995   .0739803    -4.70   0.000     -.493203   -.2027961
       south |  -.4733768   .0733317    -6.46   0.000    -.6173073   -.3294464
        west |  -.3476202     .07778    -4.47   0.000    -.5002816   -.1949588
      female |  -.0119843   .0333557    -0.36   0.719    -.0774527    .0534841
     haswork |   .0090215   .0395561     0.23   0.820    -.0686166    .0866596
       _cons |   .9952361   .1177552     8.45   0.000     .7641142    1.226358
------------------------------------------------------------------------------
```

```
. * 2SLS
. ivreg welcome (educ=padeg) age northeast midwest south west female haswork in
> come, first

First-stage regressions
-----------------------

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     33.04
       Model |  1781.14434         9  197.904927   Prob > F        =    0.0000
    Residual |  5139.67593       858  5.99029829   R-squared       =    0.2574
-------------+----------------------------------   Adj R-squared   =    0.2496
       Total |  6920.82028       867  7.98249167   Root MSE        =    2.4475

------------------------------------------------------------------------------
        educ |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0006152   .0056245     0.11   0.913    -.0104242    .0116546
   northeast |  -.0267189   .3966592    -0.07   0.946    -.8052548     .751817
     midwest |  -.4629517    .378652    -1.22   0.222    -1.206144     .280241
       south |  -.8397025   .3753322    -2.24   0.026    -1.576379   -.1030257
        west |  -.4928281   .3981003    -1.24   0.216    -1.274193    .2885363
      female |   .2527089   .1707242     1.48   0.139     -.082377    .5877948
     haswork |  -.7259447   .2024595    -3.59   0.000    -1.123319   -.3285709
      income |   .2595244   .0392934     6.60   0.000     .1824019    .3366468
       padeg |   .8834287    .074715    11.82   0.000     .7367831    1.030074
       _cons |   10.43361   .6027044    17.31   0.000     9.250659    11.61655
------------------------------------------------------------------------------


Instrumental variables (2SLS) regression

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      7.98
       Model |   20.003408         9  2.22260089   Prob > F        =    0.0000
    Residual |  192.277698       858  .224099881   R-squared       =    0.0942
-------------+----------------------------------   Adj R-squared   =    0.0847
       Total |  212.281106       867  .244845566   Root MSE        =    .47339

------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
        educ |    .034356   .0163581     2.10   0.036     .0022494    .0664626
         age |  -.0003905   .0010851    -0.36   0.719    -.0025202    .0017392
   northeast |   -.193755   .0767326    -2.53   0.012    -.3443606   -.0431495
     midwest |  -.3320943   .0737349    -4.50   0.000    -.4768162   -.1873725
       south |   -.444528   .0749109    -5.93   0.000    -.5915581   -.2974978
        west |  -.3306886   .0774515    -4.27   0.000    -.4827051   -.1786721
      female |  -.0206664   .0331846    -0.62   0.534     -.085799    .0444662
     haswork |   .0339621   .0405535     0.84   0.403    -.0456336    .1135577
      income |  -.0174456   .0091284    -1.91   0.056    -.0353621     .000471
       _cons |   .6367787   .2237823     2.85   0.005     .1975538    1.076003
------------------------------------------------------------------------------
Instrumented:  educ
Instruments:   age northeast midwest south west female haswork income padeg
------------------------------------------------------------------------------


.
. * Degree
. * Naive regression
. reg welcome degree age income northeast midwest south west female haswork

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     10.04
       Model |  20.2303516         9  2.24781684   Prob > F        =    0.0000
    Residual |  192.050754       858  .223835378   R-squared       =    0.0953
-------------+----------------------------------   Adj R-squared   =    0.0858
       Total |  212.281106       867  .244845566   Root MSE        =    .47311


------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      degree |    .069716   .0145593     4.79   0.000       .04114     .098292
         age |  -.0008963   .0010428    -0.86   0.390    -.0029431    .0011505
      income |  -.0145172   .0077197    -1.88   0.060    -.0296689    .0006346
   northeast |  -.1925321   .0766642    -2.51   0.012    -.3430034   -.0420609
     midwest |  -.3292728   .0733176    -4.49   0.000    -.4731756   -.1853701
       south |  -.4527295   .0726542    -6.23   0.000    -.5953303   -.3101287
        west |   -.322218   .0771454    -4.18   0.000    -.4736337   -.1708022
      female |  -.0188291   .0330112    -0.57   0.569    -.0836212     .045963
     haswork |   .0302061   .0393072     0.77   0.442    -.0469433    .1073555
       _cons |   .9852923   .1150745     8.56   0.000     .7594318    1.211153
------------------------------------------------------------------------------
```

```
. * Reduced form is the same as above
. * 2SLS
. ivreg welcome (degree=padeg) age northeast midwest south west female haswork
> income, first

First-stage regressions
-----------------------

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     27.74
       Model |  262.984058         9  29.2204509   Prob > F        =    0.0000
    Residual |  903.876541       858  1.05346916   R-squared       =    0.2254
-------------+----------------------------------   Adj R-squared   =    0.2173
       Total |   1166.8606       867  1.34585998   Root MSE        =    1.0264

------------------------------------------------------------------------------
      degree |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0063004   .0023587     2.67   0.008     .0016709    .0109299
   northeast |  -.0381379   .1663429    -0.23   0.819    -.3646245    .2883487
     midwest |    -.27283   .1587914    -1.72   0.086     -.584495    .0388351
       south |  -.3188269   .1573992    -2.03   0.043    -.6277594   -.0098943
        west |  -.3655764   .1669472    -2.19   0.029    -.6932492   -.0379037
      female |   .0950955   .0715948     1.33   0.184    -.0454261     .235617
     haswork |  -.2987434   .0849033    -3.52   0.000     -.465386   -.1321009
      income |   .0895094   .0164781     5.43   0.000     .0571674    .1218515
       padeg |   .3764698   .0313325    12.02   0.000     .3149725     .437967
       _cons |   .2304822   .2527499     0.91   0.362    -.2655984    .7265627
------------------------------------------------------------------------------


Instrumental variables (2SLS) regression

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      7.98
       Model |  20.1047931         9   2.2338659   Prob > F        =    0.0000
    Residual |  192.176313       858  .223981717   R-squared       =    0.0947
-------------+----------------------------------   Adj R-squared   =    0.0852
       Total |  212.281106       867  .244845566   Root MSE        =    .47327

------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
      degree |   .0806203    .038376     2.10   0.036     .0052984    .1559422
         age |  -.0008773    .001045    -0.84   0.401    -.0029284    .0011738
   northeast |  -.1915983   .0767495    -2.50   0.013     -.342237   -.0409596
     midwest |  -.3260039   .0741098    -4.40   0.000    -.4714617   -.1805461
       south |  -.4476729   .0745196    -6.01   0.000    -.5939349   -.3014109
        west |  -.3181473   .0783005    -4.06   0.000    -.4718304   -.1644643
      female |   -.019651   .0331302    -0.59   0.553    -.0846767    .0453748
     haswork |   .0331063    .040438     0.82   0.413    -.0462628    .1124753
      income |  -.0157456   .0086966    -1.81   0.071    -.0328148    .0013236
       _cons |   .9766545   .1184981     8.24   0.000     .7440743    1.209235
------------------------------------------------------------------------------
Instrumented:  degree
Instruments:   age northeast midwest south west female haswork income padeg
------------------------------------------------------------------------------


.
. * Bachelor or more
. * Naive regression
. reg welcome bachmore age income northeast midwest south west female haswork

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     10.05
       Model |  20.2506268         9  2.25006965   Prob > F        =    0.0000
    Residual |  192.030479       858  .223811747   R-squared       =    0.0954
-------------+----------------------------------   Adj R-squared   =    0.0859
       Total |  212.281106       867  .244845566   Root MSE        =    .47309


------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    bachmore |    .180604   .0376405     4.80   0.000     .1067257    .2544824
         age |  -.0010094   .0010425    -0.97   0.333    -.0030555    .0010367
      income |  -.0119683   .0076237    -1.57   0.117    -.0269315     .002995
   northeast |  -.1986287     .07665    -2.59   0.010    -.3490721   -.0481853
     midwest |  -.3295865   .0733093    -4.50   0.000    -.4734731   -.1856999
       south |  -.4597127   .0725286    -6.34   0.000    -.6020671   -.3173584
        west |  -.3223567   .0771385    -4.18   0.000     -.473759   -.1709544
      female |  -.0162441   .0329959    -0.49   0.623    -.0810062     .048518
     haswork |   .0208553   .0391608     0.53   0.594    -.0560068    .0977174
       _cons |   1.031603   .1145041     9.01   0.000     .8068622    1.256344
------------------------------------------------------------------------------
```

```
. * Reduced form is the same as above
. * 2SLS
. ivreg welcome (bachmore=padeg) age northeast midwest south west female haswor
> k income, first

First-stage regressions
-----------------------

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     18.58
       Model |  27.1072358         9  3.01191509   Prob > F        =    0.0000
    Residual |  139.086313       858  .162105259   R-squared       =    0.1631
-------------+----------------------------------   Adj R-squared   =    0.1543
       Total |  166.193548       867   .19168806   Root MSE        =    .40262

------------------------------------------------------------------------------
    bachmore |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0027875   .0009252     3.01   0.003     .0009714    .0046035
   northeast |   .0174361   .0652517     0.27   0.789    -.1106356    .1455077
     midwest |  -.1044873   .0622895    -1.68   0.094    -.2267449    .0177703
       south |  -.0892832   .0617433    -1.45   0.149    -.2104689    .0319025
        west |  -.1406104   .0654888    -2.15   0.032    -.2691473   -.0120734
      female |   .0217309   .0280847     0.77   0.439    -.0333918    .0768536
     haswork |  -.0624411   .0333052    -1.87   0.061    -.1278104    .0029281
      income |   .0212178   .0064639     3.28   0.001     .0085309    .0339047
       padeg |   .1326518   .0122909    10.79   0.000     .1085281    .1567755
       _cons |   -.148548   .0991468    -1.50   0.134    -.3431467    .0460506
------------------------------------------------------------------------------


Instrumental variables (2SLS) regression

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      7.97
       Model |  19.8836451         9   2.2092939   Prob > F        =    0.0000
    Residual |  192.397461       858  .224239465   R-squared       =    0.0937
-------------+----------------------------------   Adj R-squared   =    0.0842
       Total |  212.281106       867  .244845566   Root MSE        =    .47354

------------------------------------------------------------------------------
```

```
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    bachmore |   .2288029   .1089749     2.10   0.036     .0149142    .4426916
         age |  -.0010071   .0010435    -0.97   0.335    -.0030552     .001041
   northeast |  -.1986624   .0767232    -2.59   0.010    -.3492496   -.0480753
     midwest |  -.3240925   .0742992    -4.36   0.000    -.4699221    -.178263
       south |  -.4529486   .0740026    -6.12   0.000    -.5981959   -.3077012
        west |  -.3154482   .0785909    -4.01   0.000    -.4697012   -.1611951
      female |  -.0169564    .033062    -0.51   0.608    -.0818482    .0479353
     haswork |   .0233082   .0395421     0.59   0.556    -.0543024    .1009188
      income |   -.013384   .0082008    -1.63   0.103      -.02948    .0027119
       _cons |   1.029224   .1147246     8.97   0.000     .8040506    1.254398
------------------------------------------------------------------------------
Instrumented:  bachmore
Instruments:   age northeast midwest south west female haswork income padeg
------------------------------------------------------------------------------

.
. * Grad level or more
. reg welcome grad age income northeast midwest south west female haswork

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      9.88
       Model |  19.9371733         9  2.21524147   Prob > F        =    0.0000
    Residual |  192.343933       858  .224177078   R-squared       =    0.0939
-------------+----------------------------------   Adj R-squared   =    0.0844
       Total |  212.281106       867  .244845566   Root MSE        =    .47347

------------------------------------------------------------------------------
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        grad |   .2614784   .0562791     4.65   0.000     .1510176    .3719392
         age |  -.0016042   .0010509    -1.53   0.127    -.0036669    .0004585
      income |  -.0090316   .0075664    -1.19   0.233    -.0238825    .0058193
   northeast |  -.1661351   .0770282    -2.16   0.031    -.3173209   -.0149494
     midwest |  -.3088471   .0737815    -4.19   0.000    -.4536604   -.1640338
       south |  -.4398096   .0730472    -6.02   0.000    -.5831817   -.2964375
        west |  -.3087652   .0774797    -3.99   0.000    -.4608372   -.1566932
      female |  -.0079452   .0330403    -0.24   0.810    -.0727946    .0569041
     haswork |   .0267746   .0392807     0.68   0.496    -.0503228    .1038721
       _cons |    1.02196   .1146521     8.91   0.000     .7969287    1.246991
------------------------------------------------------------------------------
```

```
. * Reduced form is the same as above
. * 2SLS
. ivreg welcome (grad=padeg) age northeast midwest south west female haswork in
> come, first

First-stage regressions
-----------------------

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =     11.39
       Model |  7.92536569         9  .880596188   Prob > F        =    0.0000
    Residual |  66.3280905       858  .077305467   R-squared       =    0.1067
-------------+----------------------------------   Adj R-squared   =    0.0974
       Total |  74.2534562       867  .085644125   Root MSE        =    .27804

------------------------------------------------------------------------------
        grad |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0036185   .0006389     5.66   0.000      .0023644    .0048726
   northeast |   -.115661   .0450607    -2.57   0.010     -.2041032   -.0272189
     midwest |   -.153435   .0430151    -3.57   0.000     -.2378622   -.0690078
       south |  -.1482657    .042638    -3.48   0.001     -.2319527   -.0645787
        west |  -.1496585   .0452245    -3.31   0.001      -.238422   -.0608949
      female |  -.0181553   .0193944    -0.94   0.349     -.0562213    .0199107
     haswork |  -.0633957   .0229995    -2.76   0.006     -.1085376   -.0182537
      income |   .0050983   .0044638     1.14   0.254     -.0036629    .0138595
       padeg |   .0643945   .0084877     7.59   0.000      .0477354    .0810535
       _cons |  -.0251006   .0684676    -0.37   0.714     -.1594843     .109283
------------------------------------------------------------------------------


Instrumental variables (2SLS) regression

      Source |       SS           df       MS      Number of obs   =       868
-------------+----------------------------------   F(9, 858)       =      7.85
       Model |  16.8202526         9  1.86891695   Prob > F        =    0.0000
    Residual |  195.460853       858  .227809852   R-squared       =    0.0792
-------------+----------------------------------   Adj R-squared   =    0.0696
       Total |  212.281106       867  .244845566   Root MSE        =    .47729

------------------------------------------------------------------------------
```

```
     welcome |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        grad |    .471331   .2262671     2.08   0.038     .0272292    .9154328
         age |  -.0020749   .0011678    -1.78   0.076    -.0043669    .0002172
   northeast |  -.1401584   .0822475    -1.70   0.089    -.3015883    .0212716
     midwest |  -.2756809   .0820386    -3.36   0.001    -.4367008    -.114661
       south |  -.4034946   .0828198    -4.87   0.000    -.5660478   -.2409414
        west |  -.2770816   .0848178    -3.27   0.001    -.4435562   -.1106069
      female |  -.0034272   .0336392    -0.10   0.919    -.0694519    .0625975
     haswork |   .0389018   .0415716     0.94   0.350    -.0426922    .1204959
      income |  -.0109323   .0078813    -1.39   0.166    -.0264011    .0045365
       _cons |   1.007067    .116618     8.64   0.000     .7781768    1.235957
------------------------------------------------------------------------------
Instrumented:  grad
Instruments:   age northeast midwest south west female haswork income padeg
------------------------------------------------------------------------------


.
.
. log close
      name:  <unnamed>
       log:  C:\Users\qyue01\Box\final.log
  log type:  text
 closed on:  14 Dec 2018, 16:58:22
-------------------------------------------------------------------------------
```