Comparing health expenditure between foreign-born and native-born United States residents

Abstract

Objective. To investigate whether foreign-born residents (immigrants who were born outside the US but currently live in the US) have different health expenditure than people who were native-born.

Data sources. Data were taken from the 2013 Medical Expenditure Panel Survey, Household component, which is a large-scale survey of families and individuals covering topics such as health insurance, the specific health services that they use, how often they use them, the costs of these services, etc. After excluding inapplicable observations, the final size of the data set was 17,305 observations.

Study design. A two-part regression model was used to analyze this data set. Other tests were also performed to determine multicollinearity, heteroskedasticity and the appropriateness of log-transformation for the second part of the regression model.

Principal findings. This paper found that the probability of having any health expenditure is 3.3 percentage point higher for people who were born in the US (P = 0.000). Among those with some level of health spending, being born in the US was also associated with 33.7% increase in expected expenditure compared to being born outside the US (P = 0.000). These results have been adjusted for heteroskedasticity. Old immigrants had a 5.1 percentage point higher in probability of having some level of health expenditure compared to new immigrants (P = 0.000)

but there was no difference in spending between the two groups when they had some level of spending (P = 0.306).

Conclusions. Despite popular belief that immigrants (people who were born outside the US) are a burden on the system, this paper showed that they actually spent much less on health care than US-born people. Future policies should look at improving access to care for immigrants.

Introduction

From 1965 to 2015, the US population grew from 193 million to 324 million, with about 55% of that growth attributed to immigrants and their descendants (Pew Research Center, 2015). The Kaiser Family Foundation also estimated that about 13% of the current population are immigrants (Kaiser Family Foundation, 2016). Even though immigrants are an important force of the economy, the views of immigrants' impact on the US are often equally split. In 2004, 45% people surveyed by the Pew Research Center saw immigrants as a strength rather than a burden for the country whereas 44% people said that they are a burden. This gap grew ten percentage points in 2015 when 51% people saw them as a strength and 41% saw them as a burden (Pew Research Center, 2015). Despite an overall more positive attitude in the public, a greater divide was apparent.

The negative views on immigrants might affect health policies that will in turn affect the well-being of immigrants and their children. Studies have been conducted to investigate the claim that immigrants represent a financial burden on the US health care system. Goldman and colleagues (2006) and Mohanty and colleagues (2005) found that immigrants had lower health expenditure than US-born persons. Nevertheless, Palinkas and colleagues (2003) pointed out that immigrants might have different health concerns at different stages in their acclimatization to life in the United States. They often arrive with physical and psychiatric trauma, infectious diseases (Palinkas et al., 2003) but generally healthier than the general population (Wieland et al., 2016). However, the longer they reside in the country, the more susceptible they are to chronic diseases such as cancer, diabetes, hypertension and coronary heart disease (Palinkas et al., 2003; Wieland et al., 2016). For those reasons, this paper attempted to build on previous

results by using nation-wide data from the 2013 Medical Expenditure Panel Survey to examine health care expenditures of immigrants, specifically new immigrants and old immigrants, compared to US-born persons.

Conceptual framework

Health expenditure was the outcome of interest in this paper. The main variable of interest was whether the person was a US-born person, new immigrant or old immigrant. Numerous studies have pointed out that immigrants encounter significant barriers like language, cultural differences (Fowler, 1998 and Tanjasiri, 2007). This might affect their level of health expenditure when they first arrive. However, as they settle down and assimilate, they take on the American lifestyle and become more comfortable in navigating the system. This could balance out any differences they have with US-born persons. This paper included a variable on health status to control for any differences between the health status of immigrants and US-born persons. Other variables that are known to affect health expenditure that this paper also included were age, family income, insurance status, gender, region and race.

Method

The paper used a two-part model to analyze the predictors for people with $0 expenditures and people who have some expenditures. The null hypothesis was that there was no difference in health expenditure among people who were born in the US and immigrants for both parts of the model. The first part used the linear probability model and the second part used ordinary least squares regression model with log-transformed health expenditure. The analysis used log-transformed expenditure because the effect of the variables in the model was expected to yield

proportional change and not a level change. The model also used a quadratic transformation for age, as age is known to have a quadratic effect rather than a linear effect on health expenditure.

Probability of having any health expenditure = $\beta_0 + \beta_1 BORNUS + \beta_2 AGE + \beta_3 AGE^2 + \beta_4 FAMINC + \beta_5 UNINS + \beta_6 FEM + \beta_7 NORTH + \beta_8 MIDWEST + \beta_9 WEST + \beta_{10} MINORITY + \beta_{11} GOODHSTATUS$

Logged positive health expenditure = $\beta_0 + \beta_1 BORNUS + \beta_2 AGE + \beta_3 AGE^2 + \beta_4 FAMINC + \beta_5 UNINS + \beta_6 FEM + \beta_7 NORTH + \beta_8 MIDWEST + \beta_9 WEST + \beta_{10} MINORITY + \beta_{11} GOODHSTATUS$

The second test would be to stratify immigrants to compare the expenditure of new immigrants and old immigrants. The null hypothesis was that there was also no difference in expenditure for new immigrants and old immigrants.

Probability of having any health expenditure = $\beta_0 + \beta_1 BORNUS + \beta_2 OLD\_IMMIGRANT + \beta_3 AGE + \beta_4 AGE^2 + \beta_5 FAMINC + \beta_6 UNINS + \beta_7 FEM + \beta_8 NORTH + \beta_9 MIDWEST + \beta_{10} WEST + \beta_{11} MINORITY + \beta_{12} GOODHSTATUS$

Logged positive health expenditure = $\beta_0 + \beta_1 BORNUS + \beta_2 OLD\_IMMIGRANT + \beta_3 AGE + \beta_4 AGE^2 + \beta_5 FAMINC + \beta_6 UNINS + \beta_7 FEM + \beta_8 NORTH + \beta_9 MIDWEST + \beta_{10} WEST + \beta_{11} MINORITY + \beta_{12} GOODHSTATUS$

The Box-Cox test was used to test for evidence of using a log-transformed model for the second part of the model. The Variance Inflation Factors test was used to test for multicollinearity in the model. Lastly, the White test was used to test for heterosckedasticity.

Data

The analysis was done using 2013 Medical Expenditure Panel Survey (Household Component). Observations that had values other than 1 and 2 for the Born in the US variable were dropped. Observations that had values of -9, -8 or -7 for the Years in the US variable were dropped. In addition, observations that had age less than 0 or region equal to -1 or family income less than 0 were dropped. Observations that had health status less than 1 were dropped. The number of observations in the final sample was 16,659 (See Table 1).

The dependent variable was the total amount of health expenditures in 2013. The first three independent variables were whether the person was born in the US and if they were not, whether they were a new immigrant (has been in the US for less than 10 years) or an old immigrant (has been in the US for 10 years or more). Other variables are age (in years), family income, insurance status (insured or uninsured), gender (female or male), regions (north, Midwest, west or south with south being the reference group), race (white or non-white) and good health status (good or poor).

Results

The results from the two-part model showed that immigrants had substantially lower health expenditure than those who were born in the US (Table 2). Among those with no expenditures (model 1), being born in the US would increase the possibility of having expenditures by 7.8 percentage points, holding all others constant (P = 0.000). Being uninsured would decrease the possibility of having any health expenditure by 30 percentage points (P = 000). Being non-white also had the same effect, albeit much smaller, only 2.9 percentage points (P = 0.000). Being

female or living in the Midwest would increase the possibility of having health expenditures by 7.7 percentage points (P = 0.000) and 5.6 percentage points (P = 0.000), respectively. The marginal effect of age on the possibility of having health expenditures would be (-0.32 + 0.00649*2*AGE) percentage points (P = 0.000). Having good health status would decrease the probability of having any health expenditures by 12.4 percentage points (P = 0.000).

Among people with some level of health expenditures (model 2), being born in the US would be associated with a 34.5% increase (P = 0.000) in expected health expenditures. However, being uninsured would lower expected health expenditures by 89.9% (P = 0.000). Being non-white would also lower expected expenditures but only by 8.6% (P = 0.008). Being female increased expected expenditures by 18% (P = 0.000) and living in the North or Midwest increased expected expenditures by 12.8% (P = 0.001) and 15.7% (P = 0.000), respectively, compared to living in the South. When there was some expenditures, the marginal effect of age was a (2.6 + 0.00417*2*AGE) percent increase in expenditures (although P_{AGE} = 0.000 was statistically significant but P_{AGE2} = 0.124 was not). Having good health status would decrease the expected expenditures by about 74% (P = 0.000).

We rejected the null hypotheses and concluded that there was a difference in health expenditures between people who were born in the US and immigrants.

We also rejected the null hypothesis that there was no difference in probability of having any health expenditures between old immigrants and new immigrants. From table 2 model 3, old immigrants had a higher probability of having some level of health expenditures compared to new immigrants, holding all others constant (5.1 percentage points, P = 0.000). This is an effect

independent of age, that is, old immigrant had a higher probability of having expenditures not because they are older but rather, because they have been in the US longer. When immigrants had some spending (Table 2 model 4), there was no significant difference between the level of spending between new immigrants and old immigrants (P = 0.306). We failed to reject the null hypothesis that there was no difference in health expenditures between old immigrants and new immigrants when they had some level of spending.

In order to see if the data has multicollinearity, a Variance Inflation Factors test was used. The null hypothesis was that there was no multicollinearity in this dataset. The mean VIF obtained was 3.64 (Table 3). Age and Age squared had values of about 15 because age squared was a function of age so correlation was expected. We failed to reject the null hypothesis and concluded that there was no multicollinearity.

To see if a log transformation was appropriate for positive health expenditures, a Box-Cox test was conducted. The theta value obtained was 0.0058 (Table 4). The two null hypotheses were theta = 1 (A level model was the true model) and theta = 0 (A log model was the true model). With a theta value of 0.0058, we rejected the null hypothesis that theta = 1 and conclude that a level model was not the true model. We failed to reject the null hypothesis that theta = 0. We had evidence to suggest that log transformation might be appropriate for our model.

Lastly, White tests were conducted to test for heteroskedasticity for the two-part of our model. The null hypothesis is that the dataset has homoskedasticity. For the first part which used linear probability model, we expected that heteroskedasticity would be a problem. The test result for the first part had a p-value of 0.000 so we rejected the null hypothesis as expected and

concluded that there was heteroskedasticity (Table 5). For the second part, the p-value obtained was also 0.000 so we also rejected the null hypothesis and concluded that there was heteroskedasticity as well (Table 5).

To adjust for heteroskedasticity, this paper used the Feasible Generalized Least Squares approach. The new results were largely identical to the base models (Table 2, FGLS model 1 and FGLS model 2). The probability of having any expenditures decreased by about half to 3.3 percentage point for people who were born in the US (P = 0.000). When the person had some expenditures, being born in the US was associated with a slightly smaller increase in expenditures compared to the base model: 33.7% (P = 0.000). The standard errors only varied by 0.001 for the model 1s. The sign and magnitude of other coefficients and standard errors were also consistent with the base models. The only significant difference was for the West region where living in the West would now decrease the probability of having expenditure by 5.8 percentage points (P = 0.000).

Because the model has heteroskedasticity and the second part of the model used log transformation, we would need to do an adjustment to get the expected value (mean) instead of a median value from exponentiating the log value.

Using the FGLS model 1, we predicted the probability of having any health expenditures for a man who was born in the US, aged 30, with a family income of $40,000/year, uninsured, living in the Midwest, having good health status and is white. This person would have about 55.6% probability of having any health expenditures. Consider another person with the same

characteristics but was not born in the US and who is a minority. His probability of having any health expenditures would be about 46.4%.

Using the FGLS model 2, we predicted the health expenditures of a woman who was born in the US, aged 60, with a family income $45,000/year, insured, living in the Midwest, having good health status and is white. The predicted health expenditures for this person was $12,052. The predicted health expenditure for another person with the same characteristics but who was not born in the US and is a minority was $7,830.

Discussion

Contrary to the public's sentiment, immigrants had significantly less health expenditures than people who were born in the US. This is consistent with the finding of Goldman and colleagues (2006) and Mohanty and colleagues (2005). Even though the uninsured rate among immigrants was about 52%, the lower health expenditures in immigrants was independent of insurance status. This means immigrants used less health services not because they did not have health insurance. They also used less health services whether they had insurance or not. Same for family income: immigrants had lower family income but the difference was only about $3,650 (Table 1) and yet after controlling for this, immigrants still had less spending. The result was also not because immigrants had better health status than US-born persons as this had been controlled for. This finding would substantiate the qualitative research that found immigrants to have considerable barriers in access to care. The fact that old immigrants had slightly higher probability of having health expenditure than new immigrants might suggest that the longer

immigrants stay in the US, the more comfortable they are navigating the system and overcoming some of these barriers.

Instead of fearing that immigrants are a drain on resources, policy makers should understand that immigrants actually use much less health care resources because their access to care is limited. Emphasis should be placed on eliminating barriers by providing translation services, cultural competency training for health care workers, specific education outreach campaigns to immigrants and so on.

Limitations

This paper used the foreign-born variable as a proxy to determine whether the person was an immigrant or not. Because there are American citizens who were born outside of the US, the foreign-born variable might inflate the real number of immigrants. The foreign-born variable could be combined with information on whether another language other than English is spoken at home to produce a sub-set of actual immigrant population. However, doing so will leave out immigrants from countries like England or Australia where the mother tongue is also English. This is a limitation of using an existing dataset like MEPS compared to doing an individual study. There could also be other factors affecting expenditure that were not controlled for that could limit the findings of this paper.

APPENDIX

References

Goldman, Dana P., James P. Smith, and Neeraj Sood. "Immigrants and the cost of medical care." Health Affairs 25.6 (2006): 1700-1711.

Fowler N. Providing primary health care to immigrants and refugees: The north hamilton experience. CMAJ. 1998;159(4):388-391.

Kaiser Family Foundation. “Health Coverage and Care for Immigrants”. California. January, 2016.

Mohanty, Sarita A., et al. "Health care expenditures of immigrants in the United States: a nationally representative analysis." American Journal of Public Health 95.8 (2005): 1431-1438.

Pew Research Center. “Broad Public Support for Legal Status for Undocumented Immigrants”. Washington, D.C. California. June, 2015.

Pew Research Center. “Modern Immigration Wave Brings 59 Million to U.S., Driving Population Growth and Change Through 2065: Views of Immigration’s Impact on U.S. Society Mixed.” Washington, D.C.: September, 2015.

Palinkas LA, Pickwell SM, Brandstein K, et al. The journey to wellness: Stages of refugee health promotion and disease prevention. Journal of Immigrant Health. 2003;5(1):19-28.

Tanjasiri SP, Kagawa-Singer M, Foo MA, et al. Designing culturally and linguistically appropriate health interventions: The "life is precious" hmong breast cancer study. Health Educ Behav. 2007;34(1):140-153. doi: 1090198105285336 [pii].

Wieland ML, Weis JA, Hanza MM, et al. Healthy immigrant families: Participatory development and baseline characteristics of a community-based physical activity and nutrition intervention. Contemporary clinical trials. 2016;47:22-31.

Tables

Table 1: Demographics of US-born persons and immigrants, 2013

	Total	Born outside US	Born inside US	p-value
Count, (%)	16659 (100)	3503 (20.7)	13156 (79.3)	
Old immigrant, n	2675	2675	0	0.000
New immigrant, n	828	828	0	0.000
Age, mean (sd)	35.2 (22.2)	43.0 (17.1)	33.1 (22.9)	0.000
Family income (`000) mean (sd)	59.64 (56.81)	56.75 (53.64)	60.40 (57.60)	0.001
Insurance status				
Uninsured, n	2774	1198	1576	0.000
Insured, n	13885	2305	11580	0.000
Gender				
Female, n	8720	1813	6907	0.433
Male, n	7939	1690	6249	
Region				
North, n	2694	672	2022	0.000
Midwest, n	3017	363	2654	0.000
West, n	4831	1412	3419	0.000
South, n	6117	1056	5061	0.000
Race				
White, n	11066	2294	8772	
Minority, n	5593	1209	4384	0.075
Health status				
Good, n	16182	3386	12796	0.057
Poor, n	477	117	360	

Table 2: Regression results for base models, stratified models and FGLS models

	Base, model 1	Base, model 2	Stratified immigrants m3	Stratified immigrants m4	FGLS model 1	FGLS model 2
	b/se	b/se	b/se	b/se	b/se	b/se
Born in the US	0.078***	0.345***	0.117***	0.283***	0.033***	0.337***
	(0.007)	(0.036)	(0.013)	(0.071)	(0.008)	(0.036)
Age	-0.003***	0.026***	-0.003***	0.026***	-0.002***	0.026***
	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.002)
Age squared	0.000***	0.000	0.000***	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Family's income	0.000***	0.000**	0.000***	0.000**	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Uninsured	-0.301***	-0.899***	-0.299***	-0.901***	-0.318***	-0.890***
	(0.008)	(0.045)	(0.008)	(0.045)	(0.012)	(0.048)
Female	0.077***	0.180***	0.076***	0.180***	0.044***	0.162***
	(0.006)	(0.026)	(0.006)	(0.026)	(0.005)	(0.026)
North	0.011	0.128**	0.011	0.128**	-0.028***	0.124**
	(0.008)	(0.039)	(0.008)	(0.039)	(0.007)	(0.040)
Midwest	0.056***	0.157***	0.057***	0.156***	0.031***	0.152***
	(0.008)	(0.037)	(0.008)	(0.037)	(0.006)	(0.037)
West	-0.000	0.009	-0.001	0.009	-0.058***	-0.004
	(0.007)	(0.034)	(0.007)	(0.034)	(0.006)	(0.033)
Minority	-0.029***	-0.086**	-0.028***	-0.087**	-0.059***	-0.095***
	(0.006)	(0.028)	(0.006)	(0.028)	(0.006)	(0.029)
Good health status	-0.124***	-1.343***	-0.123***	-1.344***	-0.140***	-1.354***
	(0.017)	(0.074)	(0.017)	(0.074)	(0.012)	(0.070)
Old immigrant			0.051***	-0.079		
			(0.015)	(0.077)		
Constant	0.863***	6.930***	0.825***	6.992***	0.963***	6.959***
	(0.021)	(0.094)	(0.024)	(0.112)	(0.016)	(0.090)
N	16659	13464	16659	13464	15707	13464
r2	0.152	0.220	0.153	0.220	0.121	0.237

Table 3: Multicollinearity test result

Variable	VIF	1/VIF
age13x	15.00	0.066689
age2	14.72	0.067938
west	1.33	0.751115
midwest	1.24	0.804842
north	1.24	0.808641
unins13	1.17	0.851899
bornusa	1.16	0.859427
faminc13	1.07	0.937115
minority	1.03	0.970990
goodhstatus	1.03	0.973524
female	1.01	0.990981
Mean VIF	3.64	

Table 4: Cox-Box test result

Number of obs = 13,464
LR chi2(11) = 3338.49
Log likelihood = -118905.78 Prob > chi2 = 0.000

totexp13	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/theta	.0058452	.0040519	1.44	0.149	-.0020964	.0137869

Estimates of scale-variant parameters

	Coef.
Notrans	
bornusa	.3606857
age13x	.0268481
age2	.0000447
faminc13	6.68e-07
unins13	-.9343148
female	.1868723
north	.1347751
midwest	.1642302
west	.0088351
minority	-.0889884
goodhstatus	-1.406843
_cons	7.088304
/sigma	1.576364

Test H0:	Restricted log likelihood	LR statistic chi2	P-value Prob > chi2
theta = -1	-152076.58	66341.60	0.000
theta = 0	-118906.82	2.08	0.149
theta = 1	-145402.15	52992.75	0.000

Table 5: Heteroskedasticity test result

First part (Linear Probability)

```
White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(65)     =   1667.18
         Prob > chi2  =    0.0000

Cameron & Trivedi's decomposition of IM-test

---------------------------------------------------
              Source |       chi2     df      p
---------------------+-----------------------------
  Heteroskedasticity |    1667.18     65    0.0000
            Skewness |    7020.95     11    0.0000
            Kurtosis |     930.23      1    0.0000
---------------------+-----------------------------
               Total |    9618.36     77    0.0000
---------------------------------------------------
```

Second part (Least squares regression)

```
White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(65)     =    203.37
         Prob > chi2  =    0.0000

Cameron & Trivedi's decomposition of IM-test

---------------------------------------------------
              Source |       chi2     df      p
---------------------+-----------------------------
  Heteroskedasticity |     203.37     65    0.0000
            Skewness |      78.42     11    0.0000
            Kurtosis |      21.05      1    0.0000
---------------------+-----------------------------
               Total |     302.84     77    0.0000
---------------------------------------------------
```

Do file

```
*select variables
keep totexp13 bornusa yrsinus unins13 faminc13 age13x sex region13 racev1x rthlth31

*droppping observations
drop if bornusa<0
drop if yrsinus<-1
drop if age<0
drop if region==-1
drop if faminc<=0
drop if rthlth31<1

* recode variables for ease of interpretation
replace bornusa=0 if bornusa==2
gen oldimmigrant=(yrsinus>3)
gen newimmigrant=(-1<yrsinus & yrsinus<4)
gen north=(region13==1)
gen south=(region13==3)
gen midwest=(region13==2)
gen west=(region13==4)
gen female=(sex==2)
replace unins=0 if unins==2
gen minority=(race>1)
gen age2=age^2
gen goodhstatus=(rthlth31<5)

* Two-part model
gen totexp_bin=totexp13>0
gen totexp_log = log(totexp13)
* Main test: People born in the US vs immigrants
regress totexp_bin bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus /* part 1*/
estimates store m1
regress totexp_log bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0 /* part2 */
estimates store m2
* Second test of interest: People born in the US vs new immigrants vs old immigrants
regress totexp_bin bornus old age13 age2 faminc unins fem north midwest west minority
goodhstatus /* part 1*/
estimates store m3
```

```
regress totexp_log bornus old age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0 /* part2 */
estimates store m4

* Multicolinearity test
regress totexp13 bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus
vif
* Box-Cox test for log
boxcox totexp13 bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0
* Heteroskedasticity test
regress totexp_bin bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus
estat imtest, white
regress totexp_log bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0
estat imtest, white

* Remedy for heteroskedasticity
* FGLS for LPM part
regress totexp_bin bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus
predict p
sum p, detail
regress totexp_bin bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus [aw=1/(p*(1-p))]
estimates store FGLS_m1
* FGLS for log part
regress totexp_log bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0
predict e, residual
gen e2 = e^2
gen loge2=log(e2)
regress loge2 bornus age13 age2 faminc unins fem north midwest west minority goodhstatus
predict loge2hat
gen exploge2hat=exp(loge2hat)
regress totexp_log bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0 [aw=1/exploge2hat]
estimates store FGLS_m2

* Prediction
* First part, LPM
```

```
di 0.9632 + 0.0335*1 + (-0.0022)*30 + 0.0000426*(30^2) + 0.000000346*40000 + (-0.3176)*1 +
0.0440*0 + 0.0312*1 + (-0.0589)*0 + (-0.1400)*1
di 0.9632 + 0.0335*0 + (-0.0022)*30 + 0.0000426*(30^2) + 0.000000346*40000 + (-0.3176)*1 +
0.0440*0 + 0.0312*1 + (-0.0589)*1 + (-0.1400)*1
* Second part
* Predict logy
di 6.9594 + 0.3365*1 + 0.0256*60 + 0.0000429*(70^2) + 0.000000824*45000 + (-0.8898)*0 +
0.1618*1 + 0.1518*1 + (-0.0947)*0 + (-1.3537)*1
di 6.9594 + 0.3365*0 + 0.0256*60 + 0.0000429*(70^2) + 0.000000824*45000 + (-0.8898)*0 +
0.1618*1 + 0.1518*1 + (-0.0947)*1 + (-1.3537)*1
* Adjust factor
regress totexp_log bornus age13 age2 faminc unins fem north midwest west minority
goodhstatus if totexp13>0
predict resid, residual
gen resid2 = resid^2
regress resid2 bornus age13 age2 faminc unins fem north midwest west minority goodhstatus if
totexp13>0
predict s2hat
* Prediction for y
di exp(8.03909 + 0.5*s2hat)
di exp(7.60789 + 0.5*s2hat)

* Table 1
table1 bornus oldimm newimm age13 faminc unins fem north midwest west minority
goodhstatus, by(born)

* Table 2
estout m1 m2 m3 m4 FGLS_m1 FGLS_m2, cell(b(star  fmt(%5.3f)) se(par) ) stat(N r2, fmt(%3.0f
%4.3f)) label varwidth(20) modelwidth(12)
```