

# Predicting the Labor Force Participation of Female Immigrants to the U.S.

---

ALISON HARFORD

## I. Introduction

Labor market analysis is a branch of economics where decisions made at the micro level by individuals within households can have macroeconomic implications for an entire country. The history of development in the United States has shown that the labor force is a crucial factor in growth. Entry into the labor force by different groups over time has had a positive impact on the economy and the social sphere. Specifically, the immigration waves of the nineteenth and twentieth centuries transformed the United States. In addition, the female wartime workforce of World War II had a tremendous impact on the U.S. economy. In recent decades, we have witnessed the permanent entry of women into the workforce and we continue to see immigrants settling and working in the United States (Blau, et al. 2006).

The purpose of my research is to analyze the decisions of female immigrants in the United States and to predict the probability of whether or not an individual will choose to participate in the labor market. Female immigrants face unique concerns when deciding whether to enter the labor force that are not shared by male immigrants. I hypothesize that individuals from countries with high GDP per capita will be more likely to enter the labor force than their counterparts from countries with lower GDP per capita. Under the human capital theory, this is due to market specific skills such as education and training that prepare an individual for entry into the labor market. The analysis demonstrates that GDP per capita is significant in explaining labor market participation, but that it has a large negative effect.

## II. Background and Theory

The theoretical basis for this study is multifaceted. The theory of time allocation is the main framework for analyzing labor market participation. This theory indicates that individuals choose to allocate time to the labor market based on the comparison of the real wage and their reservation wage. The real wage is the wage that can potentially be earned in the labor market. This is determined by firms' labor demand and a combination of education, labor market experience, training and relevant skills on the supply side. The reservation wage is the value that the individual places on non-market time. If the individual finds that the wage rate is greater than the reservation wage, she will choose to enter the labor market. Similarly, if the individual finds that the reservation wage is greater than the wage rate, she will choose not to enter the market (Blau, et al 2006).

Additional factors influence patterns of participation. The human capital theory indicates that inputs such as education, training, skills, experience and language abilities increase employability. Employers want to hire workers who possess relevant skills so that they can minimize training costs and increase productivity. Specific to immigrants, the human capital theory indicates that the ability to communicate in the language of the host country is an important human capital input, particularly among those individuals with higher levels of education (Sandford, 2002).

The testable hypothesis is that women who immigrate to the United States from countries with high GDP per capita will be more likely to participate in the U.S. labor market than women who immigrate to the United States from countries with lower GDP. This hypothesis is

logical based on the economic theory. GDP per capita is a good way to compare countries and assess the strength of their labor markets. High GDP indicates a stronger economy and suggests that women are more likely to attain a higher level of education, obtain more skills relevant to work in an advanced market economy and have more human capital that is transportable among labor markets. Gains in human capital demonstrated by higher levels of education, skill and experience give rise to a higher potential wage rate in the U.S. market, therefore increasing the likelihood of labor market participation (Blau, et. al 2006). Individuals that possess these skills will be perceived by employers to have high levels of productivity and will therefore be in demand by employers (Garg, 2005).

### III. Literature Review

The literature reviewed prior to this analysis deal primarily with decisions by workers in the Australian labor market. Between 1975 and 1994, immigration contributed to 40% of Australia's population growth. This rapid influx of workers demonstrates that the Australian labor market is an area of interest for researchers (Wooden, 1997).

The research on labor market participation by immigrants has been led by Cobb-Clark at Australian National University. Several of Cobb-Clark's papers were of relevance to this current analysis. While her research focuses on individuals who immigrated to Australia rather than to the United States, many of the same ideas contribute to our understanding of this topic.

A 1999 Australian report shows that formal education and English language ability are positively related to labor market participation. Using the Longitudinal Survey of Immigrants to Australia (LSIA) data, Cobb-Clark and Chapman (1999) explain that these correlations reflect human capital theory and raise the market wage for individuals, therefore increasing the likelihood of participating.

A study by Cobb-Clark and Crossley (2001) tests the family investment hypothesis. The purpose of this research was to examine the effects of comparative advantage and gender. The family investment hypothesis is the idea that between married couples, one partner will work to provide support so that their spouse can develop their human capital skills so that they may be eligible for better employment opportunities. The

partner that develops their human capital is said to have the comparative advantage for labor market employment. This study also makes use of the LSIA and Cobb-Clark and Crossley (2001) find that within immigrant families that are considered "traditional," males have the comparative advantage. The hypothesis is not supported within "non-traditional" families, where the females seek to develop these skills and the males do not. These findings support human capital inputs like education and language ability as important determinants of participation. This research demonstrates that immigrant families face a unique decision when choosing how to allocate time between the labor market and non-market activities.

Cobb-Clark and Connolly (2001) examined the labor supply decisions of immigrant spouses and found that their behavior was related to the number of hours worked by the spouse, presence of children, and English language ability. Kossoudji and Cobb-Clark (2002) designed a model to determine the effect of legal status after the 1986 Immigration Reform and Control Act offered amnesty to illegal workers in the United States. They find that legal status has a positive benefit that is reflected by earnings. Though not directly related to my study of female immigrants, their research suggests that legal status may influence the labor supply decision of immigrants. Furthermore, it is important to note that the experiences of undocumented workers in "under the table" positions will not be reflected in my empirical model.

Wooden (1997) focuses on the presence of young children to determine what effect they have on the labor force participation of married women. He uses the 1991 Australian Population Census Data to estimate a multinomial logit model. The dependent variable is predicted as either full-time participation, part-time participation or no participation. A number of explanatory variables are used in this study, including age, year of arrival, English language ability, husband's income, whether other adults are present in the household, educational attainment, whether the individual is a full-time student, state of residence within Australia, the number of children in the household, and the age of the youngest child in the household.

Wooden's study divides his sample into two groups: immigrants with English speaking backgrounds (ESB) and immigrants with non-English speaking backgrounds (NESB) based on

birthplace. He admits to the imperfection of this method of classification as a short-coming of his study. Additionally, while Wooden's data does not contain an estimation of non-labor income, he uses spouse's income as a control variable and finds that there is an inverted U-shape relationship between the husband's income and the wife's labor market behavior. Initial results indicate that NESB immigrants are less likely to participate in the workforce than ESB immigrants. However, of those NESB immigrants that do work, their behavior is less responsive to the presence of children in the household than ESB immigrants (Wooden 1997). Similar to Wooden's study, several of the same control variables are included in this analysis, including age, English language ability, educational attainment and number of children present in the household.

Garg (2005) examines the effect that the interaction between gender and ethnicity has on wage differentials of female immigrants to the United States. Garg focuses on ethnicity, defined by country of origin which is organized by region, and compares results for immigrant males and females to native males and females. Her study includes several measures of human capital and is most directly related to my research. Garg confirms her hypothesis and finds that the interaction between gender and ethnicity has a statistically significant effect on wages for immigrant women.

#### **IV. Empirical Model**

The empirical model is an ordinary least squares (OLS) regression with a dichotomous dependent variable. The model aims to predict the probability of labor market participation by explaining the factors that influence the decision to participate. Focusing on GDP per capita of the country of origin as the main explanatory variable and using the remaining variables as controls is an effective way to test the hypothesis that female immigrants from countries with higher GDP per capita will be more likely to participate in the female labor market than female immigrants from countries with lower GDP per capita.

The empirical model includes the following independent variables to predict labor market participation: non-labor income, number of own children, educational attainment, age and country of origin. Country of origin is expressed as GDP per capita in that country. Non-labor

income is the sum of all income that is not earned in the labor market. This includes spouse's wages, transfer payments and other family income. Non-labor income is expected to have a negative effect, so that as non-labor income increases, the individual will work less and spend more time doing non-labor or leisure activities. The number of own children living in the household is expected to have a negative sign due to the trade-off that many women face between working and staying at home to raise their children. Childcare costs can also affect this decision. Women whose childcare costs are greater than earnings will choose to abstain from working and provide the care themselves. The age variable is expected to have a positive effect, as those who have been in the United States for a longer period of time may have more labor market experience, an expanded network of contacts and greater familiarity with the U.S.

The education variables encompass all levels of formal education. The data express education in terms of the highest level of formal education attained as of 1999. Dummy variables are defined for each category of education. The "less than high school" variable includes all individuals whose highest educational attainment is less than a high school diploma. It also includes those individuals that have no formal schooling. The "high school diploma or GED" variable includes those individuals who either obtained a high school diploma or earned a passing score on the General Educational Development (GED) Test. While a high school diploma and passing the GED do not necessarily send the same signal to employers, they indicate completion of approximately the same level of education. Since the focus of this study is not the effect of education on labor force participation and educational attainment serves as a control variable, it is appropriate to combine these respondents into one category. The "less than high school" variable is excluded from the OLS analysis so that the coefficients for the remaining education dummy variables may be interpreted as being in comparison to those that have less than a high school education.

The "some college" variable includes those who attended some college but did not earn a college degree. The "associate's degree" variable includes cases where an associate's degree was earned by completion of either an occupational or academic program. The "bachelor's degree" variable includes those cases where the highest

educational attainment was an undergraduate includes those who earned a master’s degree. Finally, the “terminal degree” variable includes those who earned either a doctoral or professional (terminal) degree. This variable includes individuals who have obtained the highest level of education available within their field of study. Each dummy education variable is expected to have a positive sign.

degree. Similarly, the “master’s degree” variable dichotomous with participation taking the value of one and not participating taking the value of zero. This study tests the hypothesis that female immigrants from countries with higher GDP per capita will be more likely to participate in the labor force. Therefore, this variable is the focus of the empirical model and the remaining independent variables serve as control variables. Table 1 is a summary of variable definitions and expected signs.

The variable that is being predicted is labor force participation. The dependent variable is

**Table 1: Variable Definitions and Expected Signs**

<b>Dependent Variable</b>	<b>Definition</b>	<b>Expected Sign</b>
LFP	Labor Force Participation rate; one if participation; zero if no participation	
<b>Explanatory Variables</b>		
GDP per Capita	Country of origin represented by GDP per capita in 2007	+
Non-Labor Income	Spouse’s income, non-labor income, other family income	-
Number of Children	The number of own children living in the household	-
Age	Age (Including only individuals 18-65 years old)	+
LessHS	Educational attainment less than HS diploma	+
HS or GED	Educational attainment equal to HS diploma or GED	+
Some College	Attended some college	+
Associate’s Degree	Educational attainment occupational or academic program	+
Bachelor’s Degree	Four year college degree	+
Master’s Degree	Educational attainment master’s degree	+
Terminal Degree	Educational attainment of doctoral or professional degree; terminal degree in field	+

The operational form is:

$$LFP = b_1 + b_2GDPcapita + b_3NONLABINC + b_4CHILDREN + b_5AGE + b_6LessHS + b_7HSorGED + b_8SOMECOLLEGE + b_9BACHELORS + b_{10}MASTERS + b_{11}TERMINALDEGREE$$

## V. Data

The data for this study come from the Current Population Survey (CPS) from March 2007 (Integrated Public Use Microdata Series, Current Population Survey). Using the CPS allows access to more recent data which will reflect more recent immigration trends. The country-specific data on GDP and population is from 2007 and comes from the World Development Indicators, which is accessible through the World Bank (World Bank). It should be noted that using GDP per capita values in 2007 is not a perfect way to reflect the characteristics of the economy in the home country, since the immigrants included in the CPS survey were already residing in the United States during that year. However, it is still a good measure since a country's relative position does not change very quickly over time.

The World Development Indicators provided information on 210 countries. Of this, 79 countries are excluded because we do not observe cases where individuals list those countries as their birthplace. In addition, the following countries are missing GDP data from the World Bank data set: American Samoa, The Bahamas, Barbados, Bermuda, Cuba, Guam, Democratic Republic of Korea (North Korea), Kuwait, Myanmar, Puerto Rico, Somalia, U.S. Virgin Islands, and Zimbabwe. Furthermore, Afghanistan was missing population data, and Iraq was missing both GDP and population information from the World Bank data set. This gives us a total of 15 countries that will be excluded from the analysis due to missing or incomplete data from the World Development Indicators. These are countries that we would have liked to include because we observe cases where individuals have these countries codes as their birthplace, but the lack of data prohibits us from doing so.

There are five more countries that are excluded due to coding problems in addition to incomplete GDP or population data. Azores is a sovereign state but it is not included in the World Development Indicators data, therefore there is no GDP or population information available. Czechoslovakia split into the Czech Republic and Slovakia. We observe cases where individuals listed one of these countries as their birthplace, but there is no definitive way to code cases with this country listed as the birthplace, since there is separate data available for both the Czech

Republic and Slovakia. In order to keep the analysis as accurate as possible, these cases are excluded. Yugoslavia is listed as a birthplace, but this country no longer exists, as it was broken up into several other countries. Similar to the Czechoslovakia problem, we observe cases where individuals were born in one of the new sovereign states and therefore, do not know how to properly code individuals who list Yugoslavia as their birthplace. Kosovo is recognized as a sovereign state by the United States, but this acknowledgement is not universal. As a result, there is no available data on GDP/population. Taiwan is the final country excluded from the analysis; since the World Bank does not classify it as a separate nation and there is also no available data on GDP or population.

The 2007 CPS data contains 125,888 cases. Of these, 103,426 have the U.S. listed as their birthplace. This leaves us with 22,462 individuals born outside of the United States. There are a total of 3365 cases excluded due to the lack of GDP or population data, and the coding problems described above. Since we are examining labor force participation, the data set is filtered to only include women between the ages of 18 and 65 born outside of the United States.

## VI. Results

This empirical model is statistically significant at the .001 level. The adjusted R square value is .045. This value is very low, and indicates that this model fails to explain all of the variance. Additionally, while each explanatory variable is statistically significant at the .001 level, three of the explanatory variables have the wrong sign. GDP per capita was predicted to have a positive effect, whereas the actual effect is negative. Number of children was predicted to have a negative effect but we find that the real effect on labor market participation is small but positive. The age variable also has the wrong predicted sign. It was expected to have a positive effect, but the regression analysis indicates a negative effect. From these results we conclude that for immigrant women, country of origin measured by GDP per capita has a negative impact on labor market participation. Due to the fact that the effect of this variable is negative, we do not find support for the original hypothesis. Refer to Table 2 on the following page for the results of the analysis.

---

**Table 2: OLS Regression Results**

Variable	Coefficient (B)
GDP per Capita	-4.869E-7 (-5.669)**
Non-Labor Income	-8.870E-7 (-15.394)**
Number of Children	.019 (19.016)**
Age	-.002 (-22.071)**
LessHS	(excluded for comparison)
HS or GED	.159 (40.992)**
Some College	.155 (36.998)**
Associate's Degree	.242 (47.717)**
Bachelor's Degree	.244 (57.391)**
Master's Degree	.281 (49.667)**
Terminal Degree	.334 (41.287)**

\*\*Indicates statistical significance at the .01 level  
(t-statistics in parentheses)

Adjusted R Square=.045

---

One possible explanation for the negative effect of GDP per capita might be that women from countries with higher GDP per capita maintain a higher reservation wage that determines entry into the labor market. Another explanation might be the failure to control for the number of hours worked. The purpose of using GDP per capita as a measure of country of origin was a way to correct for disparities between nations, but this analysis failed to distinguish between full-time and part-time employment for individuals. We might improve on this model by taking this into account and controlling for the number of hours worked per week.

The empirical model employed in this analysis has several shortcomings. The most notable is the absence of English language ability as a control variable. Previous literature suggests that English language ability is a significant determinant of labor market participation, income, and the type of work that is attained (managerial or lower skilled positions). Unfortunately, the CPS data did not include a measure of English language ability or skill.

Continuing this project with a different data set might correct for this deficiency. Secondly, inclusion of additional control variables might improve upon this model. Future research might include control variables that better reflect the characteristics of the home country, such as proximity to the United States and whether the home country is an English speaking nation.

Finally, an OLS regression might not be the best way to predict a dichotomous dependent variable. Ordinary least squares estimates the dependent variable and creates a line of best fit. When estimating labor market participation, the dependent variable can only assume two values – participate or not participate. As a result of this, the dependent variable cannot be less than zero or greater than one but the OLS analysis can make predictions outside of the zero to one range. The literature indicates that a probit model is a more appropriate way to predict a dependent variable like this, and future research might take advantage of this technique.

## VII. Conclusion

The results of this multivariate analysis indicate that GDP per capita is a significant factor in determining the labor market participation of female immigrants to the United States. Though the results do not confirm the hypothesis that female immigrants from countries with higher GDP per capita will be more likely to participate in the labor force than female immigrants from countries with lower GDP per capita, this study does add to the literature on this subject by laying a foundation for future research in this area. As stated earlier, there are several ways to improve upon the empirical model by adding additional control variables or estimating the model using a different data set. Additionally, using a probit model instead of an OLS regression may be a more appropriate way to analyze the labor market participation of female immigrants to the United States. At this time, this analysis does not suggest any policy implications, as further research must be conducted in order to more accurately determine the factors influencing labor force participation.

## REFERENCES

- Blau, Francine D., Marianne A. Ferber, and Anne E. Winkler.** "The Allocation of Time between the Household and the Labor Market." *The Economics of Women, Men and Work*. 5th ed., Upper Saddle River: Pearson/Prentice Hall, 2006, pp. 85-135.
- Cobb-Clark, Deborah, and Bruce J. Chapman.** "The Changing Pattern of Immigrants' Labor Market Experience." *Australian National University Centre for Economic Policy Research*. Discussion Paper No. 396, 1999.
- Cobb-Clark, Deborah and Marie D Connolly.** "A Family Affair: The Labor Market Experience of Immigrant Spouses." *Social Science Quarterly*, 2001, 82(4).
- Cobb-Clark, Deborah, and Thomas F. Crossley.** "Gender, Comparative Advantage and Labour Market Activity in Immigrant Families." *Australian National University*. Discussion Papers No. 433, 2001.
- Garg, Mahi.** "Wage Differentials for Immigrant Women in the United States: The Heightened Effect of Gender and Ethnic Interaction." *Undergraduate Economic Review Illinois Wesleyan University*, 2005, 1.
- King, Miriam, et al.** *Integrated Public Use Microdata Series, Current Population Survey: Version 2.0*. [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004. <cps.ipums.org/cps>.
- Kossoudji, Sherrie A., and Deborah A. Cobb-Clark.** "Coming Out of the Shadows: Learning about Legal Status and Wages from the Legalized Population." *Journal of Labor Economics*, 2002, 20(3).
- Sandford, Jeremy.** "English Language Proficiency and the Earnings of Mexican Immigrants." *Illinois Wesleyan University: Senior Honors Project*, 2002.
- Wooden, Mark.** "Family Composition and the Labor Supply Choices of Married Immigrant Women." *Australian Journal of Labor Economics*, 1997, 1(2), pp. 121-42.
- World Bank.** *World Development Indicators, 2007*. <<http://go.worldbank.org/IW6ZUUHUZO>>.