2018

No Hablo Inglés: Language and Skill Upgrading in Mexican Migration

Paul Lapinski
Skidmore College, plapinsk@skidmore.edu

No hablo inglés:

Language and Skill Upgrading in Mexican Migration

By

Paul Lapinski

A Thesis Submitted to

Department of Economics

Skidmore College

In Partial Fulfillment of the Requirement for the B.A Degree

Thesis Advisor: Qi Ge

May 1, 2018

Abstract

Much of the literature on international migration finds a brain drain – the most skilled migrants leave their home country and stay abroad in an economy where they can maximize their performance. Studies of return migration, however, suggest the possibility of a brain gain; home countries can be made more productive if migrants engage in skill upgrading while abroad and bring their abilities back. Using data from the Mexican Migrant Project I estimate the return to migration, focusing on the role that English plays as a skill in the labor market. To overcome the possibility of endogeneity via self-selection I utilize an Instrumental Variable method, using who a migrant stays with as the instrument. Estimates show that being a migrant increases Mexican wages by 36.4% but that English has a negligible effect on economic outcomes.

Acknowledgements

I would like to thank my thesis (and pre-thesis) advisor, Professor Qi Ge, for providing consistent support and important feedback over the academic year, helping me to manage the magnitude of this project. I would also like to acknowledge and thank Jiebei Luo for her help in managing the preliminary steps of my data analysis, which was crucial to my success. Special thanks go to my research partner and advisor Professor Smriti Tiwari for playing a part in the inspiration for this project as well as for teaching me almost everything I know about using Stata.

I. Introduction

One of the most valuable tools that an individual in the labor market has at his disposal is his human capital. A worker has the ability to increase his labor outcomes by becoming more educated, gaining on-the-job experience, and learning any skills that make him a more desirable asset. In developing countries, one of the more popular ways to engage in such human capital accumulation is through international migration. Many times, the economies of foreign countries offer opportunities that cannot be experienced in an individual's home country, or participating in an activity abroad is more beneficial than participating in that same activity while at home. When a potential migrant finds either of these to be the case when assessing his labor market opportunities, they he choose to immigrate (Dustmann et al., 2011).

Since 1960, the United States has attracted the most immigrants of any other country in the world – almost one fifth of all migrants reside there. According to statistics gathered by the Migration Policy Institute, among the more than 43.7 million immigrants in the U.S in 2016, more than a quarter of them were from Mexico (11.6 million) (Batalova, Hallock, and Zong 2018). Although their rate of entry has leveled off in the last decade, Mexican immigrants continue to lead as the largest group of immigrants in the U.S as they have since the 1970s. Because of their significant presence abroad, it is important to examine how Mexican immigrants are interacting within the U.S labor market and how their experiences there affect economic outcomes later in life. This paper focuses on the high volume of return migration that occurs between the U.S and Mexico – many Mexican migrants travel to the U.S for the primary purpose of earning money to take back home after a relatively short period of time. While doing so they are exposed to new industries and production techniques that may contribute to their human

capital. By understanding how migration can help and hurt individuals we can construct policy measures that lead to the most mutually beneficial outcomes for all parties involved.

Using data from the Mexican Migrant Project, I pursue a study of the returns to migration, focusing on how English language abilities serve as a skill to upgrade human capital. To estimate whether or not being a migrant and having superior English skills leads to a wage premium in the Mexican labor market, I use both an Ordinary Least Square (OLS) and Two Stage Least Squares (2SLS) model, controlling for demographic characteristics, community effects, and certain migration elements. With the OLS model, I find that return migrants experience a significant 36.4% increase in Mexican wages compared to those who are not migrants. For those who can speak and understand English, wages increase by 18.4%. However, due to the possible endogeneity present in the OLS model, I also test the effect of English skills with a 2SLS model, using who a migrant stays with in the U.S as an instrument. My main specification indicates that English skills are an insignificant determinant of wages. The paper proceeds as follows: section II reviews relevant literature on worker productivity, return migration, and language as a skill; section III discusses my data and empirical methodology; section IV presents my results and discusses their implications; section V conducts a robustness check; and section VI concludes.

II. Literature Review

Productivity

While a migrant may intend to engage in skill upgrading while he is abroad, it is possible that the actual trip is spent participating in activities that do not actively contribute to increasing worker productivity. Indeed, there are many reasons why an individual may choose to migrate, – to send money back home to his family, to avoid political unrest at home, to enter a new job –

however, no matter the intention, migrating always has the potential to be an investment in human capital by increasing productivity (Sjaastad, 1962). A migrant may, for example, move to another country with plans to work in a comparatively prosperous agricultural sector in order to earn higher wages and bring money back home to his family, but while working he is exposed to technology that is scarcely available in his country of origin. Exposure to this technology may provide the migrant with an advantage in his job back home or even with the ability to upgrade into a higher-income occupation back home.

The first theory of productivity in relation to migration comes from the Roy (1951) model, which focuses on the transferability of skills to different occupations. In making the decision of whether or not to move, migrants compare potential earnings at home and abroad, a calculation that rests on the return to the skills that they possess. If the rate of return to skills is higher in the destination, higher-skilled individuals will migrate, whereas if the rate of return is higher at home, lower-skilled workers will migrate. Borjas (1987) adapts the Roy model and Sjaastad's (1962) theory of migration, providing a theoretical framework that makes three conclusions about the emigration rate: (1) an increase in mean income in the home country decreases the emigration rate; (2) an increase in mean income in the U.S increases the emigration

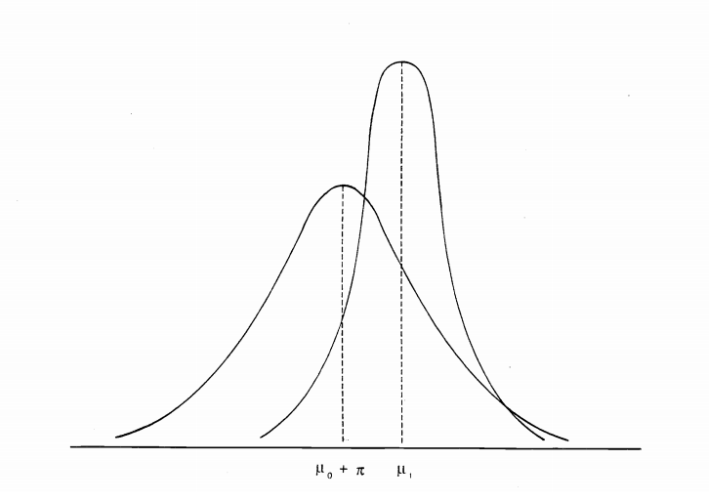

Figure 1
Emigration rate distribution conditional upon mean income
(Borjas, 1987)

rate; (3) an increase in the costs to migration decreases the emigration rate. If income distribution varies in the U.S more than it does at home, high-income workers will move away and low-income workers will stay, causing a brain drain to take place in the sending country (Borjas 1987). These positively selected immigrants continue on to eventually earn more than the native population. But, as income inequality increases in the home country, more low-skilled workers will want to immigrate and high-skilled workers will stay home; the former group takes advantage of the economy in which higher earners are being taxed more than lower earners. Figure 1 is taken from Borjas's (1987) seminal paper to demonstrate the scale effect that he describes. If $\mu_0 + \pi$ represents mean incomes in the home country plus mobility costs and μ_1 is mean incomes in the potential destination country, it is clear that as $\mu_0 + \pi$ increases, the emigration rate increases, and the quality of migrants decreases; lower-income individuals are moving to get away from the high-income inequality at home. But as $\mu_1 > \mu_0 + \pi$, emigration rates of the lower-skilled drop off and the emigration rate decreases.

Return Migration

While migrants can increase the value of their human capital while abroad, it may be that their new skills are better applied at home than in the host country to which they traveled to learn them. In such cases, migrants may choose to return to their home country after a certain period of time. To define the situations in which return migration may occur, Dustmann et al. (2011) employ a dynamic Roy model that considers migrants having not only one but two skills, allowing comparative advantage to play a role in their decision making. In this case, migration is conditional upon relative skill as it is valued in the home versus the host country. The two-skill model leads to a conclusion that makes Borjas's (1987) model of selectivity more robust: there

may be some stayers who are more skilled than the movers and there may be some movers who are more skilled than the stayers.

Of the three iterations in their model, Dustmann et al.'s (2011) primary theory for return migration is defined as strong transferability, where experience acquired in the host country augments human capital in the host country by less than it does in the home country. Or, experience acquired in the host country augments human capital in the home country by more than experience acquired in the home country[1]. In both situations, return migration occurs. This model explains what can be called a "brain gain," the result of which brings skilled migrants back to their home country to raise average productivity as opposed to the drain (Borjas, 1987), which describes migrants that stay away after moving.

Using data from the Mexican Migrant Project (MMP) Reinhold and Thom (2013) provide an extensive review of the returns to migration in the U.S for Mexican migrants, finding that an additional year of migration experience in the U.S increases labor outcomes back in Mexico by 2.4%. I also use data from the MMP in my analysis and will describe it in more detail later in the paper, but for now offer that it is a cross-sectional database of households that are sampled in Mexican communities that are randomly selected every year. Their model is in the form of a simple OLS, but various independent variables are progressively added to observe any possible bias that may present itself. Controlling for age, education, marriage, and migration experience causes insignificant variation in this estimate, from which the authors postulate that endogeneity bias is not strongly present. While it is true that additional controls can help to alleviate the extent of omitted variable bias, it is not certain that doing so accounts for all endogeneity issues.

[1] Dustmann et al. (2011) also define partial transferability (individuals either never migrate or migrate and never return) and super transferability (some potential migrants are forced to stay at home because of certain, usually legal, barriers).

There is still the possibility of reverse causality, for example – does being a migrant increase labor market outcomes or do increased labor market outcomes encourage migration? Accounting for such empirical issues is a primary objective in my own study.

Beyond demonstrating that migrating to the U.S can in and of itself improve the outcomes of Mexican migrants at home, Reinhold and Thom (2013) also find that the return to migration experience is twice as high as the return to age in the Mexican labor market. Put simply, spending a year abroad and then returning is twice as effective in improving income as spending a year at home. Although the results are revealing and significant, the authors do not provide detailed reasoning as to why these numbers appear. They offer some brief suggestions, such as the possibility that skill upgrading in the U.S is more effective because of their developed firms' access to technological resources, but cannot make definite conclusions with the broad analysis they have conducted.

While it is fairly certain that time in the U.S can improve the labor market outcomes of migrants when they return, there is more doubt as to which kinds of individuals can benefit the most from such trips. Intuitively, it makes sense that less-skilled workers could reap the greatest rewards. Starting at lower levels in terms of education and job experience, lower-skilled workers have the most to gain from the variety of opportunities that international migration offers. Empirical studies show both that returns for lower-educated workers are smaller at home than in the U.S and that higher-educated workers benefit more from staying home. Using an interaction term between education and years of migration experience, Reinhold and Thom (2013) find that an extra year of migration and education is negatively correlated with labor outcomes in Mexico. Greenwood and Zahniser (1998) reach a similar conclusion; their results indicate that for those who have 12 or more years of education, a year in Mexico provides a return that is almost twice

that of U.S experience. The population of migrants that are highly educated may perform better by staying at home rather than ever traveling abroad because that is where their skills are most relevant and can be applied the most effectively (Winters et al., 2001). So, while low-skilled workers may not perform well while abroad (Borjas 1987), a positive return to migration for these individuals may be observed upon returning home

Migrant Performance

Part of the extent to which return migration benefits labor market outcomes depends on the occupations that a migrant holds while abroad and at home. Job-switching is both a motivation for and a consequence of migration (Sjaastad, 1962; Greenwood and Zahniser, 1998). Improved labor market outcomes following a trip abroad may be indicated by a change into a higher-income occupation. However, this is not the only marker of success. Reinhold and Thom (2013) suggest that the relationship between migration experience and Mexican earnings may be a function of Mexican-relevant skills that are learned abroad, acquired confidence and motivation that can increase productivity back home, or the quality signal that being a migrant sends to employers.

By identifying which occupations migrants work in while abroad and comparing them to jobs worked back home, Reinhold and Thom (2013) show that a year of relevant migration experience increases earnings back home by 4.6%– relevant experience is defined as years worked in a job in the U.S that is similar to a job worked back in Mexico. Similarly, Greenwood and Zahniser (1998) find that if the primary occupation a migrant works in while in the U.S matches their primary occupation in Mexico, there is a higher probability that occupation-specific human capital is transferred back to Mexico. Individuals whose occupations in Mexico and the U.S do match enjoy a 17.7% return to U.S experience, due to the long period of relevant

skill upgrading that they go through (Greenwood and Zahniser, 1998). These results beg the question of why skill-upgrading abroad is more effective than in Mexico; it may be because the U.S's more developed firms have access to more technological resources or that they are more organized, but there is no definite conclusion to explain the phenomenon. Despite the apparent benefit of staying in the same occupation, job mismatching may in fact be a source of return migration. Workers tend to stay in jobs that they are productive in and leave those in which they are not productive. When an individual has this low structural dependence on their job, they are more likely to experience job turnover (Jovanovic, 1979). If migrants from Mexico go to the U.S and work in jobs purely for the purpose of accumulating earnings to bring back home, it is more likely that they will have a lower dependence on that job, leading them to be less productive and eventually return home. The MMP allows me to identify which occupations migrants worked in during their last trip to the U.S and their job in Mexico at the time of the survey. From this data, I can construct a measure of relevant skills to make a conclusion as to its importance for return Mexican migrants.

While some of the literature on migration between Mexico and the U.S finds a mainly positive return to foreign human capital accumulation (Greenwood and Zahniser, 1998; Reinhold and Thom, 2013; Wahba, 2015), the aforementioned importance of job-matching indicates that migration may not always provide a positive outcome. Abarcar (2016) conducts a fake resume experiment in the Phillippines and finds that callback rates are 12% smaller for those with more migration experience. Employers seems to associate more time spent out of the country with a depreciation in human specific capital. Additionally, Li (2017) finds that migrants only experience a wage premium upon returning home to Mexico if they have a minimum five-year uninterrupted stay in the U.S. This length of stay is required to experience skill upgrading that

does not atrophy due to constant back-and-forth journeys between home and the destination country. Only if the rates of return to human capital are high, the skills are transferrable, and these outweigh the negative effects of labor market interruptions, will individuals experience a wage premium back home.

Li (2017) employs an Instrumental Variable (IV) method in analyzing data from the MMP, using U.S border policy changes as her instrument. She hypothesizes that rational migrants will delay planned travel (either to a foreign country or to return home) in order to avoid periods of stricter border enforcement. Using border policy changes as an instrument makes sense, as their occurrence cannot be predicted by migrants, making it a completely exogenous influence on the migration decision. However, the assumption that migrants act rationally is strong, because most migrants do travel across the border illegally. Even if we allow this assumption, it is not certain that migrants take the law into account, that it is enforced strictly where they are crossing, or, if they did plan for the policy change, whether it was a significant factor in their migration plan. In fact, migrants may delay travel simply because human capital accumulation abroad is more valuable than it is at home (Dustmann et at., 2011). Li's (2017) reasoning for using an IV is sound, however, and leaves room for other istruments to be tested in a model fo return migration.

Using a unique survey dataset of Mexican households personally collected by economists Duran and Massey, Lindstrom (1996) writes extensively about how the migration trip duration decision is made. Lindstrom (1996) proposes that the economic opportunities that are available at the point of origin have a strong influence on how long a migrant stays abroad. The length of time that is spent in the destination:

(1) Increases with lower destination wages

(2) Increases with higher migration costs

(3) Increases with higher origin wages

(4) Increases with higher origin investment opportunities.

Point (4) is an especially interesting hypothesis, based on the assumption that migrants will want to stay in a destination country longer if they know that upon returning home there will be lucrative opportunities in which to invest whatever was earned while abroad. Using a hazard model, Lindstrom (1996) finds both points (3) and (4) to be supported by the data. These results make an interesting contribution to the migration story, but it is important to recognize how wages and investment opportunities are measured in this study: using the percentage of earners with wages double that of the minimum and the percentage of women economically active, respectively. Both measures are unique and nuanced but rely on strong assumptions about the characteristics of the Mexican communities that they come from. The conclusion that arises from the analysis is attractive but may be the result of unobserved endogeneity. However, Lindstrom's (1996) results do serve as a logical compliment to and in support Li's (2017) conclusions about the length of migration trips. Based on the evidence that more migration experience is correlated with negative outcomes, I expect that individuals who have spent more time in the U.S to see smaller increase in wages.

Migration Networks

A key determinant of migration performance, and what will serve as a bridge into my discussion of the role of language, are migration networks. Migration networks play a significant role in a household's decision to send a migrant and how many to send; they also have the potential to enhance labor market outcomes. Networks have the ability to contribute directly to a migrant's experience through providing housing or financial support, as well as informational

support. The latter can inform potential migrants of the gains they can expect from migrating, therefore becoming a key economic factor in the overall decision (Winters et al., 2001). Winters et al. (2001) make a distinction between family and community networks, finding that they serve as substitutes in the migration decision. If a household has access to a strong family network, the importance of a community network becomes less significant.

Because migration networks serve an important role in determining the outcomes of migrating, it is important to account for them in empirical studies. Adsera and Pytlikova (2015) create a migration network measure by calculating the foreign population of the sending country that is living in the host country as a fraction of the total foreign population. While this certainly captures part of the migration story, it is a crude measure. It is likely that not all migrants from the same country know each other, or that they even migrate to the same areas. The data from the MMP that I will be employing helps to alleviate this problem by indicating with who the Mexican migrants live with in the U.S. I am able to identify whether a migrant lives with a relative, a random community member, or by themselves. While these connections certainly give some individuals an advantage as a migrant, it is possible that they also serve to retard the rate of human capital accumulation a migrant experiences while abroad. I elaborate on this in the next section.

Language as a Factor in Migration

There seems to be uncertainty as to what combination of migration factors leads to which labor market outcomes – agreement on the optimal length of migration, level of education, and occupation to maximize returns has yet to be reached. Studying migrants' abilities with the destination country's language, specifically English, may help to fill in some of the gaps that exist in the migration story. Adsera and Pytlikova (2015) suggest that linguistic proximity may

play a significant role in a potential migrant's decision. Additionally, recent studies find that knowledge of a "widely spoken" language (such as English) may help to improve migrants' economic outcomes; this is an especially appealing prospect to the return migrants who travel abroad for a short time to acquire skills before returning home to apply these skills to jobs that pay higher wages.

Two measures of language that are commonly used to determine linguistic distance are the Levenshtein distance and the Dyen index. The Levenshtein measure relies on phonetic dissimilarity of words in two languages, computing the number of steps that it takes to move from a given word in one language to the other. The Dyen index is based on the similarity of samples of words from two languages. Based upon these indices, countries that share a similar language experience a 14-20% increase in emigration rates (Adersa and Pytlikova, 2015). This effect may be larger between countries that are geographical neighbors, that share historical pasts, or have a greater disparity in unemployment rates. In the case of Mexico and the U.S, all of these traits are present. Not only have the two struggled to find mutual content in immigration policy in the past, but the conversation is ongoing to this day. Despite this potential for increased motivation to migrate, it is also possible that pre-migration exposure to English decreases the relevance of linguistic distance for migration to countries that are English-speaking (Adersa and Pytlikova, 2015). If that is the case, repeat migrants between the U.S and Mexico may have less incentive to continue migrating if they do not feel a need to continue improving their English skills (assuming learning a language is a significant factor for the migrant).

As with the return to migration on a broader scale, there is a question as to which kinds of migrants can benefit the most from learning a new language while abroad. It may be that knowledge of a widely spoken language is less relevant for migrants with lower average skills

(Adsera and Pytlikova, 2015). This could be because simply learning a new skill does not compensate for other areas where the individual may lack in skill. Conversely, it has already been shown that individuals with higher levels of education are less likely to migrate (Winters et al., 2001; Reinhold and Thom, 2013; Greenwood and Zahniser, 1998). If the majority of migrants are relatively low skilled, then it is logical to assume that they can benefit the most from migration, including the return to language.

Language and Skill Upgrading

While the language literature and return migration literature are robust, it seems that there lacks a connection between the two, despite the weight that one has on the other. Reinhold and Thom (2013) do identify two modes by which migrants might improve labor market outcomes with English-language ability. Firstly, if migrants learn English as an actual skill while abroad, they may be more suited for and perform better in occupations that require interactions with English-speaking foreigners. Another benefit of higher English proficiency is the extent to which migrants can learn new skills while in the U.S, because they are more suited to understand the instruction of their native employers. Beyond offering these suggestions, Reinhold and Thom (2013) do not delve deeply into the empirical effects that language has on economic success.

Adersa and Pytlikova (2015) conclude that emigration flows between two countries are larger the closer their languages are to each other. Additionally, the presence of a large network of immigrants from the same origin alleviates the pressure to learn the destination language. With my own study, I am primarily interested in discovering how migrants fair if this pressure is indeed alleviated: if a migrant does not engage in learning a new language while abroad, are their labor market outcomes back in Mexico diminished? Under Dustmann et al.'s (2011) definition of strong transferability, because the primary language in the U.S is English, it is reasonable to

assume that Mexican migrants are frequently exposed to English and therefore benefit from learning a new language. The experience of learning English in the U.S augments human capital in Mexico by more than the experience of learning English in Mexico. However, should migrants move to areas in the U.S where they are not frequently exposed to English, the assumptions of this model would be violated and the argument in favor of choosing to migrate should be reconsidered. Migrants may not experience a high level of skill upgrading because of the network that they are a part of, which will likely lead to low levels of human capital accumulation.

Hypotheses

Based on Dustmann's model of migration and human capital as well as Adersa and Pytlikova's (2015) insight into the role of language as it relates migration, I hypothesize that migrants that travel to the U.S and stay with a family member will not experience a high level of English language learning, which will lead to decreased labor market outcomes when they return home. I do not expect the magnitude of the negative estimate to be very large, as English is only one of many skills that migrants can learn. Developing a causal relationship between language and wages is difficult, due to the potential self-selection of more highly skilled migrants who speak English into higher-earning jobs. To resolve this empirical complexity, I employ an Instrumental Variable (IV) regression using who a migrant stays with as an instrument. A migrant's lodging decision may be predictable based on their available migration network, but it is less likely to be a predictor of what kind of job they have and therefore of their wage level. The use of an IV to study return migration is not an entirely new approach in the literature, however its application to language skills is, as far as I have seen, the first of its kind. By

implementing this method, I hope to bridge the gap between separate studies of language and migration.

While language and its relation to migration will be the main focus of my empirical results, I also contribute my own findings as to the labor market effects of being a return migrant versus not being a migrant. Based on Reinhold and Thom's (2013) and Greenwood and Zahniser's (1998) discovery of a wage premium for return migrants using data from the same source as myself, I predict that my results will show migrants to earn around 15% more in Mexico than non-migrants, having generated enough human capital to elevate them to a higher income bracket. Additionally, although there is ample evidence indicating that lower-skilled migrants benefit the most from return migration (Winters et al., 2001; Reinhold and Thom, 2013; Greenwood and Zahniser, 1998), I hypothesize that in terms of English skills, higher-skilled migrants will benefit the most, suggesting that it is more relevant to the kinds of jobs that they work in.

Even if it becomes apparent from my IV results that migration networks contribute to a lack of language learning, it may be that this has no significant effect on the labor market outcomes of return migrants, as their occupation does not demand knowledge of English. It will be important to identify which occupations migrants work in, if they change industries after returning to Mexico, and which skills each industry demands. By restricting my models to certain industries and occupations, I expect to see a variety of outcomes.

III. Data and Methodology

Data

I. *Mexican Migrant Project*

The data used for this study come from the Mexican Migrant Project (MMP), a longitudinal survey project by a collaboration between Princeton University and the University of Guadalajara. Since 1982, the MMP has annually sampled households randomly throughout different communities in Mexico, gathering information on families' immigrations to the United States, as well as other social, demographic, and economic characteristics. The MMP is unique in that it provides a robust snapshot of a migrant's experience while abroad alongside equally detailed information about their current status in Mexico. Because the survey sampling is random and does not collect data exclusively from migrants, it is possible to develop a causal relationship that explains the difference, if there is one, between migrants and non-migrants. This characteristic of the data is essential to my analysis.

The MMP consists of seven core files, but for my analysis I will only be utilizing two. The first is a cross-sectional database for each member of every household surveyed, whether they were living in the house at the time of the survey or not, and the second is only for heads of households that migrated to the U.S. Both files are used so that, after determining the differences (and similarities) between migrants and non-migrants, a more thorough investigation can be done concerning the factors that specifically affect migrants' outcomes when they return to Mexico.

Each individual reports their income (both in Mexico and the US) by a different unit (i.e. hourly, weekly, yearly) but I convert all values into monthly increments for my analysis, assuming a 40-hour work week for those who are paid by the hour. This is a strong assumption to make, as especially in the U.S it is possible that hourly workers work either much more or much less than 40-hours a week; however, because the majority of migrants in the data are paid on a monthly schedule, I do not expect the overall results to be significantly skewed. I also trim the data by dropping the top one percent of the income distribution to account for extreme

outliers and reporting errors in the data. Finally, I restrict the sample to include only household heads. Dropping all other individuals eliminates around 142,000 observations, leaving 9,583, but I do so because household heads are typically the oldest family members and have the most migration experience. Were this a study focusing on the relationship between migration and an entire household's productivity I would include other family members, but for the purposes of examining individual-level outcomes, looking at household heads makes the most sense.

II. *Descriptive Statistics*

Table 1 shows the summary statistics for the sample after the alterations are made. The data set contains 9,583 household heads, 12.7% of which are return migrants. I define "return migrant" as anyone who reports that they have taken a "trip" to the U.S per the MMP survey. It is unclear as to exactly how long a trip is; as the table indicates, even those who are deemed non-migrants have worked in the U.S to a certain extent. Whether these statistics are the result of the subjective interpretation of what a migration trip is or simply an error in the data, their small magnitude will have an insignificant effect on my analysis. Those who are classified as migrants have taken an average of 2.14 trips to the U.S. Also notable is that non-migrants earn more in Mexico than migrants do, $188.4 versus $177.9 USD, respectively. A two-way t-test indicates that this difference is significant to the 5% level. This disparity in wages is one of the primary motivators for my analysis, as some migration theory indicates that moving abroad increases labor market outcomes (Dustmann et al., 2011; Reinhold and Thom, 2013) while others propose that it is detrimental (Li, 2017; Abarcar, 2016). Another t-test shows that those with a greater level of English skills earn significantly more than those who are less skilled. Furthermore, Table 1 shows that migrants earn much more when working in the U.S than they do while working in

Table 1 – Summary Statistics, Mean and (Standard Deviation)			
	(i)	(ii)	(iii)
Variables	*Whole Sample*	*Non-migrants*	*Migrants*
Migrant	0.127	-	-
	(0.333)		
English	0.0277	-	0.217
	(0.164)		(0.412)
Male	0.828	0.812	0.931
	(0.378)	(0.390)	(0.253)
Age	47.74	48.18	44.78
	(15.12)	(15.22)	(14.09)
Years of Education	6.812	6.866	6.437
	(4.458)	(4.536)	(3.862)
Wage Mexico (USD)	187.1	188.5	177.9
	(163.3)	(165.2)	(149.0)
Wage US Last Trip (USD)	137.3	5.358	1,041
	(441.6)	(77.28)	(744.5)
# U.S Trips	0.273	-	2.140
	(1.100)		(2.346)
# U.S Months	6.905	-	54.19
	(30.43)		(68.63)
Job Switch	-	-	0.830
			(0.375)
Stayed with Relative	0.0491	-	0.386
	(0.216)		(0.487)
Relevant Experience	0.0311	0.00203	0.230
	(0.174)	(0.045)	(0.421)
Observations	9,583	8,362	1,221

Mexico. A thorough investigation of the data can possibly provide further insight into why such differences exist.

Other variables of interest include years of education, job switch, and relevant experience. In this sample, migrants have only about an average of half a year less of education than non-migrants, but a t-test shows this difference to be significant to the 1%. While abroad, 38.6% of migrants stayed with a relative and when they returned 83% entered an occupation different from the one they had in the U.S. Only 23% worked in a U.S occupation of the same industry category as in Mexico.

Methodology

Summary statistics indicate that there are differences between migrants and non-migrants as well as a variety of outcomes among those who are return migrants. To investigate both of these relationships I employ an Ordinary Least Squares (OLS) regression model and an Instrumental Variable (IV) regression.

I. *OLS*

To determine if there is a causal relationship between an individual's wage in Mexico and whether or not that individual is a migrant, I define an OLS model of the following log-linear form:

$$lnmxwage_i = \alpha_0 + \alpha_1 mig_i + \alpha_2 characteristics_i + \alpha_3 community_i + \varepsilon_i \qquad (1)$$

where $lnmxwage_i$ is the log monthly wage of individual *i* in USD, mig_i is a dummy variable indicating if an individual is a migrant or not, $characteristics_i$ is a vector of individual characteristics (i.e age and education), $community_i$ is a vector that controls for the community that each migrant lives in, and ε_i is an additive stochastic error term. Measuring wage as a log variable allows me to report the results as a percent increase rather than a dollar value, which may be a more appropriate fit for the data and also better represents the elasticity of wages.

As a precursor to my IV regression, I also include English proficiency as an independent variable and restrict the sample to return migrants:

$$lnmxwage_i = \alpha_0 + \alpha_1 eng_i + \alpha_2 characteristics_i + \alpha_3 community_i + \varepsilon_i \qquad (2)$$

where eng_i is a dummy variable for English proficiency. In the original data, language skills are measured by a categorical variable where: 0=Do not speak, do not understand (30.79% of sample); 1= Do not speak, but understand some (30.06%); 2= Do not speak, but understand much (17.44%); 3= Speak and understand some (17.12%); 4= Speak and understand much

(4.59%). For the purposes of my regression, I recode eng to equal 1 for categories 3 and 4, and equal to 0 for categories 1 and 2. As such, my model focuses on the ability of migrants to at least speak English. This qualification slightly limits the explain-ability of the model, but makes the empirical analysis more plausible. The results of this regression will indicate to what extent a migrant's abilities in speaking English affect their labor market outcomes when they return to Mexico. It may be that when individuals travel to Mexico a large part of their skill upgrading process involves learning English, a skill that could open up lucrative new job opportunities at home. If so, the results of the regression will show the percent increase in Mexican wages for those who possess higher level English skills.

II. *IV*

It is possible that there is a high level of endogeneity present in the OLS model, rendering it biased. Even without knowing in which direction the relationship between English skills and wages goes, there could be an element of self-selection that biases the regression coefficients. That is, it may be that those who have a higher (lower) level of English proficiency are the kinds of people who enter into occupations that benefit (do not benefit) from that skill, providing higher (lower) wages than other jobs. If this is the case, then in equation (2), eng_i would be correlated with unobservable variation in the error term, ε_i and its coefficient would be inaccurate as to the true effect.

To alleviate the effect that endogeneity has on the regression estimations, I instrument using information on who a migrant stays with when they travel to the U.S. As Winters et al. (2001) explain, migration networks play an important role in a migrant's experience abroad, especially for those who have established family connections in the U.S. While these networks certainly aid in making assimilation into the U.S more manageable, it is also possible that living

with a family member has the unintended consequence of eliminating the potential to learn English as a skill, as migrants could be speaking the language they are most comfortable with instead of practicing English skills. Because who a migrant stays with is not a predictor of wages back home, it is an appropriately exogenous relationship to leverage in the IV model.

An important caveat to the first stage of my IV regression is that the dependent variable is binary, therefore indicating that a linear model is an inappropriate fit to provide estimations. This is problematic as the first stage of a two-stage least squares (2SLS) model is linear. In order to accommodate my dependent variable, I replicate the procedure used by Adams et al. (2009), which, when applied to my data, will consist of three steps: (1) estimate a probit model of *eng* on *relative*, (2) compute the fitted probabilities of *eng* ($\hat{e}$), and (3) estimate *lnmxwage* with an IV regression using $\hat{e}$ as the instrument. Using this method accounts for the binary nature of the dependent variable, ensuring that the first stage is correctly specified; specification errors that typically appear in 2SLS models are also solved by using this extended method (Adams et al., 2009).

As opposed to using a typical probit model, I use the Stata command "dprobit," which reports the marginal effects of the probit equation. The marginal effects indicate the change in the probability for an infinitesimal change in each continuous independent variable as well as the change in the probability for a dummy variable (StataCorp, 2015). The equation is as follows:

$$eng_i = \Phi\,(\alpha_0 + \alpha_1 relative_i + \alpha_2 jobswitchl_i + \alpha_3 relexp_i + \alpha_4 USexp_i + \alpha_5 UStrips_i + \alpha_6 characteristics_i + \alpha_7 community_i + \alpha_8 occupation_i + \alpha_9 migration_i + \nu_i) \quad (3)$$

where $relative_i$ is a dummy variable indicating whether or not a migrant stays with a relative while abroad. Similar to the English skills variable, the lodging information is originally coded as a categorical variable where: 1=fellow home-community member; 2=friend; 3=employer;

4=relative; 6= did not need it/no one; 7=other. I convert each category into a dummy variable and focus on the one for relatives, hypothesizing that it will have the most significant effect on language skills (additionally, according to the dataset it is the most common accommodation for Mexicans to utilize). In addition to the variable of interest, $jobswitchl_i$ is a binary variable indicating whether or not a migrant worked in an occupation different from their job in the U.S upon returning home, $relexp_i$ is a dummy for whether not a migrant worked in the same industry category before and after migrating, $USexp_i$ is a continuous variable that measures an individual's collective experience in the U.S in months, and $UStrips_i$ measures the collective number of trips that an individual has taken in their lifetime. I also include $occupation_i$ to control for the job that a migrant works in and a vector of the additional dummies that indicate who a migrant stayed with in the U.S, $migration_i$.

Following the dprobit, I compute the fitted probabilities of $eng\,(\hat{e})$:

$$\hat{e} = Pr(eng_i = x), \tag{4}$$

where x is the different possible values of eng. Finally, using $\hat{e}$ as the instrument, the 2SLS is as follows:

Stage I:

$$eng_i = \alpha_0 + \alpha_1\hat{e} + \alpha_2 jobswitchl_i + \alpha_3 relexp_i + \alpha_4 USexp_i + \alpha_5 UStrips_i +$$
$$\alpha_6 characteristics_i + \alpha_7 community_i + \alpha_8 occupation_i + \alpha_9 migration_i + \nu_i \tag{5}$$

Stage II:

$$lnmxwage_i = \beta_0 + \beta_1\widehat{eng}_i + \beta_2 jobswitchl_i + \beta_3 relexp_i + \beta_4 USexp_i + \beta_5 UStrips_i +$$
$$\beta_6 characteristics_i + \beta_7 community_i + \alpha_8 occupation_i + \alpha_9 migration_i + \varepsilon_i \tag{6}$$

The instrument I use is logical as it meets the criteria for a valid regression instrument. If a standard IV takes the form:

$$D_i = \alpha_0 + \alpha_1 Z_i + \nu_i \quad (7)$$

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i, \quad (8)$$

where Y_i is the regressand of interest, D_i is the regressor of interest, and Z_i is the instrument, it must be that

$$E[Z_i \cdot \varepsilon_i] = 0, E[Z_i \cdot \nu_i] = 0. \quad (9)$$

In other words, the instrument must be uncorrelated with the error terms in both equations, a condition that ensures that any effect of Z on Y must be a result of the effect of Z on D (Angrist et al., 1996). In my analysis, the use of the IV will allow me to conclude that English proficiency increases or decreases a migrant's wages in Mexico because of the positive or negative effect that living with a relative has. It is also required that

$$\text{cov}(D_i, Z_i) \neq 0, \quad (10)$$

otherwise interpreted as α_1 not being equal to zero. As will be demonstrated below in the results section, my instrument, *relative*, returns a significant non-negative value, therefore fulfilling the second IV criterion. It is possible that *relative* is correlated with the error term in either (5) or (6). This may be because the relatives that migrants have could form a network that helps them get a higher paying job in Mexico. However, it is not certain that these networks are correlated with jobs for which migrants have accumulated relevant human capital. That is, relatives might provide migrants with access to a job that offers a higher wage, but the migrant did not necessarily work in a similar job while abroad and may not be aptly prepared to perform well in said occupation. Finally, for an IV to be valid, it is also required that the instrument is not correlated between individuals and that it is randomly assigned (Angrist et al., 1996). It is probable that who a migrant stays in the U.S does not affect who another migrant stays with. However, it is possible that there still exists self-selection in this model, because a migrant's

lodging choice may not be randomly assigned. Depending on a migrant's existing network, some choices may simply not be available to some individuals. I attempt to account for this possibility by including dummy variables for all of the lodging options referenced in the survey data. I also interpret a Cragg-Wald F-statistic and Anderson-Rubin Wald test for my IV regressions in the results section to test the strength of my instrument. For the most part, and with the use of key control variables, *relative* is an arguably plausible instrument, but it is important to note that the regression restrictions serve as strong assumptions in the model and will be considered when the results are interpreted.

III. *Heterogeneous Effects*

In addition to my base OLS and IV models, I also regress wages on English ability within specific occupations and industries. While it makes sense that English as an aspect of human capital can improve labor market outcomes, it is plausible that it serves no purpose at all, or even makes a negative contribution if it is not a skill that is important to jobs that Mexicans work in their home country. Therefore, I identify certain jobs in which English may have an important role, such as those in the tourism industry, sales positions, or academic positions. I then restrict (1), (2), (5), and (6) to observations that fit into these categories and compare the results to those in the overall sample.

IV. Results

I. *OLS*

Table 2 reports the estimates of log Mexican wages using an OLS regression. Columns (i) and (iii) contain the estimates for equations (1) and (2), respectively, and (ii) and (iv) include additional controls that I will use later in the IV regression, including: relevant experience in the U.S, number of trips to and months in the U.S, and vectors for occupation and who a migrant

stayed with while abroad. With limited controls in (ii), the dummy for being a migrant estimates an 11.7% increase in wages, but is insignificant. Aside from the interaction between migrant

Table 2 – OLS Estimates of Log Mexican Wages

Variables	(i)	(ii)	(iii)	(iv)
Migrant	0.117	0.364***	-	-
	(0.147)	(0.137)		
English	-	-	0.0207	0.184***
			(0.0726)	(0.0624)
Male	1.434***	0.438***	1.434***	0.436***
	(0.0290)	(0.0331)	(0.0290)	(0.0331)
Years of Education	0.0667***	0.0397***	0.0662***	0.0386***
	(0.00294)	(0.00292)	(0.00288)	(0.00288)
Age	0.0579***	0.0367***	0.0576***	0.0359***
	(0.00403)	(0.00343)	(0.00400)	(0.00341)
Age^2	-0.000704***	-0.000422***	-0.000703***	-0.000418***
	(3.90e-05)	(3.34e-05)	(3.89e-05)	(3.34e-05)
Migrant X Edu	-0.0286***	-0.0226***	-0.0242***	-0.0149**
	(0.00882)	(0.00742)	(0.00657)	(0.00597)
Migrant X Age	0.00109	-0.00168	0.00283***	0.00301**
	(0.00243)	(0.00208)	(0.00102)	(0.00121)
Relevant Experience	-	0.0191	-	0.0457
		(0.0580)		(0.0575)
# U.S Months	-	-0.00187***	-	-0.00215***
		(0.000379)		(0.000391)
# U.S Trips	-	-0.0326***	-	-0.0262**
		(0.0117)		(0.0114)
Comm. Effects	Yes	Yes	Yes	Yes
Migration Effects	-	Yes	-	Yes
Occup. Effects	-	Yes	-	Yes
Observations	9,583	9,583	9,583	9,583
R-squared	0.464	0.644	0.464	0.644

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

status and age, all of the other variables in this specification are significant to the 1% level (male, education, age, age-squared, and an interaction between migrant status and education). It is notable that the coefficients on education and its interaction with migrant status are of the opposite sign. Although those who are more educated in general perform better in the labor market, it seems that for migrants, education is detrimental, causing a 2.86% decrease in wages. This may be because migrants typically work in occupations that are more physically intensive and spending time to go to school takes the place of time that could be spent having more relevant on-the-job experience.

A similar story is shown in column (iii) for the effect of English skills on wages. With the same controls, OLS estimates an insignificant increase in wages of 2.07%. However, all of the control variables are again significant to the 1% level and indicate the same effects in terms of education and migrant status. I also note that older individuals fair better in the labor market, however outcomes go down slightly as age increases.

In specifications (ii) and (iv) I include additional controls and find that the coefficients on the migration and English dummies become highly significant and greatly increase in magnitude. It appears that the occupation and migration controls help to eliminate some of the bias that exists in the OLS model. However, it is still possible that endogeneity exists, so I now move on to the IV estimates.

II.a. *IV – Estimates*

The results of the initial dprobit regression of English skills on who a migrant stays with indicate that those who live with a relative are more likely to have higher level language skills. The marginal effects show increase in probability of 23.9%, 22.3%, 0.18%, or 0.06% for each specification in Table 3, respectively. These estimates are contrary to what I predicted in my

hypothesis, as I believed that living with a relative would make it less likely that a migrant speaks and understands English. It is possible that the estimates predict a positive causal relationship because the relatives that migrants stay with have been living in the U.S for a substantial amount of time and are adept at speaking English. As such, they help migrants to improve their language skills, serving as a human capital upgrader. They may even be better

Table 3 – IV Estimates of Log Mexican Wages

Variables	(i)	(ii)	(iii)	(iv)
English	0.334	-0.489***	-0.721***	-0.156
	(0.271)	(0.177)	(0.208)	(0.147)
Age	-	0.0596***	0.0602***	0.0368***
		(0.00447)	(0.00448)	(0.00396)
Age^2	-	-0.000726***	-0.000729***	-0.000426***
		(4.37e-05)	(4.38e-05)	(3.89e-05)
Male	-	1.412***	1.402***	0.365***
		(0.0331)	(0.0331)	(0.0393)
Years of Education	-	0.0651***	0.0655***	0.0413***
		(0.00317)	(0.00320)	(0.00327)
Job Switch	-	-	0.342***	0.197***
			(0.0658)	(0.0542)
Relevant Experience	-	-	0.165**	0.193***
			(0.0729)	(0.0639)
# U.S Months	-	-	-0.000907*	-0.00188***
			(0.000541)	(0.000428)
# U.S Trips	-	-	-0.0365***	-0.0381***
			(0.0141)	(0.0120)
Comm. Effects	-	Yes	Yes	Yes
Migration Effects	-	-	Yes	Yes
Occup. Effects	-	-	-	Yes
Observations	9,583	7,265	7,265	6,824
R-squared	0.001	0.447	0.446	0.631

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

teachers than a migrant's coworkers or U.S community members because of the increased social comfort that exists from the familial connection.

I include the first stage results of the IV regression in Table 7 in the Appendix. The coefficient for the fitted probability variable in every specification is significant to the 1% level and indicates that staying with a relative makes it highly probable that a migrant will have increased English skills. Table 3 shows the second stage results of the IV regression, using the fitted probabilities from the dprobit regression of English skills on who a migrant stays with in the U.S as the instrument. Before including all of the variables that I specify in equation (6), I run the IV with only controls for individual characteristics and community effects – the results are in column (ii). All of the variables are significant to the 1% level, and the coefficient on the main regressor indicates that being able to speak and understand English decreases Mexican wages by 48.9%, given that they lived with a relative while in the U.S. While the other independent variables return values that go in the same direction as the OLS estimates and are of similar magnitude, the English dummy indicates a relationship opposite to that shown in columns (iii) and (iv) of Table 2. If the instrument is valid, then it appears that the IV model has accounted for the unobserved variation that is present in the OLS model, revealing a more accurate relationship between English skills and Mexican wages. As I add additional controls for who a migrant stays with in (iii) and for different occupations in (iv), all coefficients maintain the same sign and all stay significant to the 1% level, except for my independent variable of interest. While specification (iv) estimates a 15.6% decrease in Mexican wages, it is an insignificant value. However, it does seem that this is the most appropriate specification in which to test the role of language in determining labor market outcomes, as the kind of job that an individual has plays a key role in how important English skills are.

Also of note in (iv) are the coefficients for the number of cumulative months a migrant has stayed in the U.S for and the total number of trips they have taken. Both are negative and indicate that more time spent in the U.S has a negative effect on wages, especially in terms of overall trips taken, which estimates a 3.81% decrease. As Li (2017) suggests, this could be because an extended period of time spent away from a migrant's home economy interrupts human capital accumulation. The more frequently a migrant moves between Mexico and the U.S, the less often they can work on becoming more productive in their job. This comes with the caveat that even if a migrant does take many migration trips, they still have the potential to increase wages if they work in the same industry both at home and abroad; the coefficient on relevant experience indicates an 19.3% increase in wages, holding all other variables constant. At the same time, if a migrant switches their occupation after returning home, they experience a 19.7% increase in wages. This seemingly contradictory estimate comes about because the job switching variable does not make a distinction between moves across jobs and across industries. In other words, a migrant can enter a new job but still have acquired relevant experience in their industry.

II.b. *IV – Instrument Validity*

Even if instruments are exogenous, they may still be considered "weak" in the sense that they are not strongly related to the dependent variable they are trying to explain via the endogenous regressor. Stock and Yogo (2002) generate a matrix of critical values to be compared against a Cragg-Donald F-statistic, which will indicate whether or not the instrument being used in a regression is weak. Both the matrix and F-statistic can be generated in STATA when the ivreg2 package is installed. The null hypothesis for this test is that the instruments are weak (the alternative being that they are not). For all four of my IV specifications, the Cragg-

Donald F-statistic is well above 1000, which easily supersedes the rule of thumb of 10 to reject the null hypothesis (Stock and Yogo, 2002) and accept that the instrument for whether or not a migrant stays with a relative is not weak.

For equations where there is only one endogenous variable present, this test is sufficient; however, if there are two or more, a Cragg-Donald may not be appropriate. In this case, an Anderson-Rubin (AR) test is used instead as it has an unbounded confidence region (Anderson and Rubin, 1949). The AR Wald test statistics are significant to the 1% for specifications (ii) and (iii) but insignificant for (i) and (iv), which makes sense alongside the coefficient estimates for the latter two equations also being insignificant. The AR test results should be interpreted with caution however, as they can indicate either that the instrument is endogenous or that β_0 is false; there is no method to identify which one (Anderson and Rubin, 1949). Because the Cragg-Donald F-statistic is undoubtedly significant, and is more commonly used when there is only one main endogenous variable as in my model, I conclude that the instrument I use is strong and significant.

III. *Heterogeneous Effects*

Tables 8 and 9 are in the Appendix and contain the estimates for equations (1) and (2) restricted to 5 different job categories: educators, manufacturing supervisors, sales employees, skilled manufacturers, and agricultural workers. The first three categories are selected because they are more likely to require English skills, being jobs that involve a degree of leadership, possible interaction with U.S firms, or with visiting tourists. The last two categories are selected for the opposite reason: it is more likely these workers will have no need to speak English while on the job. Table 8 indicates that only for those who work in sales, being a migrant increases wages by 91.7% in Mexico, significant to the 1% level. Also significant at this level, supervisors

who gain relevant experience while abroad experience a 522.5% decrease in wages. This counterintuitive result may appear because of unobservable differences in the way that firms operate in the U.S and Mexico. Perhaps managerial practices are different enough to the extent that the seemingly "relevant" experience gained by working in the same occupation both at home and abroad actually harms the migrant's human capital accumulation by teaching them techniques that do not function well in Mexico.

Specification (v) in Table 9 produces a result contradictory to my hypothesis that those who are in lower skilled occupations will benefit less from having English skills; workers in the agricultural sector experience a 26.2% increase in wages if they can speak and understand English, significant to the 1% level. On the other hand, supervisors' wages in Mexico decrease by 42.1% if they have a higher level of English speaking skills.

When broken down by industry, the results of the IV, equations (5) and (6), are insignificant in every category as can be seen in Table 10 in the Appendix. Evidently, English skills do not matter more in one industry compared to another. Notable from Table 10 is that education is significant for every industry category; a more educated individual experiences an increase in wages. The largest increase is for educators, whose wages go up by 4.41% for every year of education they have. Additionally, the coefficient on job switching is only significant for supervisors and those who work in sales, causing a 34.8% and 22.5% decrease, respectively. This suggests that relevant experience gained abroad is most important for these jobs.

Discussion

While the main focus of this paper is to examine how English skills affect labor market outcomes, I also consider what the estimates that focus solely on being a return migrant or not indicate. According to my preferred OLS specification, being a migrant has a fairly large positive

impact on labor market outcomes in Mexico, taking into account demographic differences, overall U.S experience, and differences in occupations. This result is consistent with some of the other literature on return migration (Greenwood and Zahniser, 1998; Reinhold and Thom, 2013; Wahba, 2015), but I do not consider it to be a significant contribution because of the possible flaws in the empirical model that I use. Despite the various controls that I do include, it is still likely that there is unobserved endogeneity in the form of selection bias. Eliminating this bias is a great empirical hurdle; determining an influence of the migration decision that is not highly correlated with labor market outcomes is difficult because of the intrinsic characteristics of migrants.

Although OLS estimates indicate that being skilled in speaking and understanding English has a significant and positive effect on labor market outcomes, this model does not capture the full story. Even when including a control for who a migrant stays with, OLS does not account for its relationship with English language skills. That is, I do not observe whether living with a relative helps or hurts the development of that particular skill. Employing an IV allows me to develop a valid causal relationship between skill upgrading and labor market outcomes.

Before controlling for occupations, my IV results show that the effect of learning English on Mexican wages is negative. From this I infer that it is not a skill that is valued in the Mexican labor market. In fact, my empirical estimates suggest it is detrimental to have this skill, and will not only *not* improve outcomes, but it will diminish them. This is likely a result of English learning taking the place of skill upgrading that is more relevant to that migrant's particular occupation back in Mexico. Li (2017) and Wahba's (2015) explanations for why migrants do not experience a wage premium may apply here, as they propose that frequent migration trips interrupt skill upgrading, leading to an overall elimination of any wage premium from migrating

and returning. In a similar way, even taking one trip abroad and staying with a relative may deter the development of skills that are more important in the Mexican labor market. Additionally, Wahba (2015) suggests that migrants may have a negative wage premium if they have to downgrade their occupation when going abroad. This explanation is in accordance with Jovanavic's (1979) theory of job mismatching and turnover.

Another explanation for why some of my estimates indicate a decrease in wages due to learning English is provided by Abarcar (2016). Employers may perceive migration experience to be a detrimental addition to human capital, assuming that the time spent abroad has taken away from valuable relevant experience. Even though my model and others control for relevant experience, employers have certainly not explored economic literature deeply and may favor non-migrants over migrants based on their intuition. That those who are employed as supervisors see a significant decrease in wages with English skills especially supports this explanation. In an occupation that requires overseeing and directing other people, it is logical that employers prefer someone who has been consistently working in the field.

When I do include controls for occupation, the coefficient on English skills becomes insignificant, although the sign remains negative. If this is the most accurate model to describe the causal relationship, I can still conclude that increased English skills do not matter back in Mexico but no longer have evidence that they are detrimental to wages. This conclusion is neither a rejection nor a failure of rejection of my original hypothesis. I expected that living with a relative would be detrimental to a Mexican migrant's English skills thereby decreasing their labor market outcomes back home. While the results actually show that housing with a family member improves English skills, its contribution to human capital is overall negligible on the job market.

Although I have shown that the instrument I use is a valid one, the IV model is not the unanimously agreed-upon method for studying international migration; it is clear from the literature that many studies employ OLS regressions. Both techniques have advantages and disadvantages that cause estimates from each of them to be consistently different. In a review of studies that estimate the return to schooling, Card (2001) finds that IV estimates are on average 20% larger than those from OLS, a somewhat unexpected finding considering that OLS models are typically upward biased. Two possible explanations for this result could be that measurement errors cause OLS to be downward biased or that unobserved differences between the treatment and comparison groups in the IV model cause the IV to be even more upward biased (Card 2001). Because an IV is modeled in two stages, it can be likened to using grouped data, and it may be that grouping reduces the variance in the independent variable by more than it reduces the covariance of the independent variable with the bias terms (Card 2001), thus causing an upward bias. This is not to say that an IV method does not make sense in this study. In fact, it might be best to determine what model should be used on a particular set of data based on previous literature that studies a similar concentration. For example, studies of the returns to education consistently find that IV estimates are larger than OLS, indicating it to possibly be the more logical model in that case. Not many studies of Mexican migration use an instrument, although one that I do review for this paper (Li, 2017), finds a negative return using immigration law changes as an exogenous shock. My main results also indicate a negative causal relationship, however a more profound reading of the literature is necessary to make an informed recommendation on the most useful empirical model for migration.

V. Replicability – Latin American Migrant Project

I. *Data*

Because so many immigrants come to the U.S from Mexico, it is important to study migration between the two countries so that policymakers can continue to determine the best ways to maximize their role in the labor market both abroad and at home. However, Mexico is not necessarily a representative example of all international return migration, so it is useful to apply the same empirical models to other countries to observe if and how the effects differ. Perhaps English skills matter more in countries that do not border the U.S or that receive less tourists. An appropriate region with which to replicate my original models is Latin America to the south of Mexico; countries in this area share cultural and economic similarities but their difference in geographic positioning suggests that the migration narrative may be different.

As an extension of the MMP and conducted by the same organization, the Latin American Migrant Project (LAMP) implements a similar survey as the MMP does in 11 more countries: Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Guatemala, Haiti, Nicaragua, Paraguay, Perú, and Puerto Rico. I apply equations (1) – (6) to 10 of these countries (Paraguay is excluded because of its incongruent file format with the other files) to test my original hypotheses.

II. *Descriptive Statistics*

Table 4 contains summary statistics for the LAMP data. Of the 7,566 household heads in the sample, 7.1% are return migrants, which is equivalent to 534 individuals. Unlike the migrants from Mexico, those from the other Latin American (L.A) countries are more educated than non-migrants are. These groups also differ in that the Mexican migrants on average have spent half as many cumulative months in the U.S as other L.A migrants – 54.19 versus 108.6. Also, more

migrants from L.A (58.4%) stay with relatives while they are abroad than Mexican migrants (38.6%).

The most notable difference between the two datasets is that non-migrants from L.A countries earn less at home than migrants: $224.7 USD versus $313.3 USD, while the opposite is true for the MMP data. A two-way t-test shows this difference to be significant. Additionally,

Table 4 – LAMP Summary Statistics, Mean and (Standard Deviation)

Variables	(i) *Whole Sample*	(ii) *Non-migrants*	(iii) *Migrants*
Migrant	0.071	-	-
	(0.256)		
English	0.0354	-	0.500
	(0.185)		(0.500)
Male	0.700	0.696	0.755
	(0.458)	(0.460)	(0.431)
Age	49.62	49.52	50.83
	(15.49)	(15.58)	(14.23)
Years of Education	8.402	8.369	8.837
	(4.930)	(4.966)	(4.402)
Wage	230.9	224.7	313.3
	(321.6)	(314.8)	(391.2)
# U.S Months	7.696	0.0299	108.6
	(120.6)	(2.504)	(442.1)
# U.S Trips	0.0961	0.00185	1.337
	(0.448)	(0.155)	(0.934)
Job Switch	-	-	0.891
			(0.311)
Stayed with Relative	0.0412	-	0.584
	(0.199)		(0.493)
Relevant Experience	0.0159	0.00341	0.180
	(0.125)	(0.0583)	(0.384)
Observations	7,566	7,032	534

half of the migrants from L.A are proficient English speakers and another t-test indicates that these individuals earn significantly more than those who are less proficient.

III. *Results & Discussion*

Table 5 utilizes the same four specifications as my initial OLS model and additionally controls for country effects, as there are 10 different ones represented in the data. The estimates

Table 5 – OLS Estimates, Log Latin American Wages

Variables	(i)	(ii)	(iii)	(iv)
Migrant	0.832**	1.218***	-	-
	(0.324)	(0.352)		
English	-	-	-0.176	-0.219*
			(0.138)	(0.128)
Male	1.475***	0.598***	1.477***	0.603***
	(0.0413)	(0.0491)	(0.0413)	(0.0491)
Years of Education	0.0936***	0.0530***	0.0916***	0.0507***
	(0.00449)	(0.00488)	(0.00444)	(0.00485)
Age	0.0769***	0.0584***	0.0766***	0.0577***
	(0.00698)	(0.00646)	(0.00698)	(0.00647)
Age^2	-0.000916***	-0.000665***	-0.000921***	-0.000666***
	(6.62e-05)	(6.21e-05)	(6.62e-05)	(6.22e-05)
Migrant X Edu	-0.0424***	-0.0270*	-0.0133	0.00421
	(0.0164)	(0.0149)	(0.0139)	(0.0131)
Migrant X Age	-0.0124**	-0.0206***	-0.000236	-0.00783**
	(0.00510)	(0.00472)	(0.00262)	(0.00309)
Relevant Experience	-	-0.0801	-	0.00817
		(0.159)		(0.157)
# U.S Months	-	-1.20e-05	-	-6.03e-05
		(0.000150)		(0.000149)
# U.S Trips	-	-0.170***	-	-0.110*
		(0.0599)		(0.0590)
Comm. Effects	Yes	Yes	Yes	Yes
Migration Effects	-	Yes	-	Yes
Occup. Effects	-	Yes	-	Yes
Observations	7,566	7,566	7,566	7,566
R-squared	0.391	0.521	0.390	0.520

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in columns (i) and (ii) are similar to those for the MMP data, with greater magnitudes. Being a return migrant from L.A provides a huge advantage in the labor market, increasing wages by 121.8% after controlling for community, migration network, and occupation effects. Columns (iii) and (iv) are more interesting, indicating that English skills are a disadvantage in L.A labor markets, decreasing wages by 21.9%, all else held constant; the OLS model for MMP data estimated an 18.4% increase. It may be that proximity to the U.S/Mexico border is indeed important to determining wages, however the same endogeneity issues could be present, so I turn to the IV model.

The first stage of the IV (not shown in this paper but available upon request from the author) indicates that staying with a relative while in the U.S has a positive and significant effect on the probability that a migrant has higher level English speaking skills. Like Mexican migrants, it appears that the relatives of migrants from L.A do not rely solely on Spanish to communicate and are skilled enough to speak and help improve English. As with my previous second-stage IV results, the coefficient on the English variable is insignificant in specification (iv), leaving me unable to demonstrate causality in the relationship between English skills and home country wages. That the LAMP shares this insignificance with the MMP data supports my main conclusion that English language skills have an overall negligible influence on labor market outcomes for return migrants. This could be the result of a combination of the factors that I detailed in my earlier discussion. Like employers in Mexico, employers in L.A could perceive migrants to be less productive due to their migration trips interrupting the accumulation of human capital. At the same time, a high demand for jobs that require a higher level of English skills may cancel out the negative effects, overall generating a neutral effect on wages.

However, as with the OLS model, it is notable that the sign on the English coefficient is opposite that in my original estimations – it is negative in the MMP data and positive in the LAMP. Evidently, there is some difference between the role that English plays in Mexico and the rest of L.A. Perhaps there are more tourists that travel to L.A, as these countries may be more appealing as vacation destinations compared to Mexico. If that is the case, then being a better English speaker could potentially provide a migrant with a wage premium in their home

Table 6 – IV Estimates, Log Latin American Wages

Variables	(i)	(ii)	(iii)	(iv)
English	0.516**	-0.271	-0.0738	0.148
	(0.231)	(0.198)	(0.431)	(0.303)
Age	-	0.0768***	0.0774***	0.0526***
		(0.00702)	(0.00701)	(0.00714)
Age^2	-	-0.000923***	-0.000928***	-0.000629***
		(6.67e-05)	(6.66e-05)	(6.80e-05)
Male	-	1.443***	1.444***	0.705***
		(0.0417)	(0.0417)	(0.0530)
Years of Education	-	0.0895***	0.0895***	0.0559***
		(0.00439)	(0.00439)	(0.00516)
Job Switch	-	-	0.00583	-0.0429
			(0.190)	(0.158)
Relevant Experience	-	-	-0.235	0.00458
			(0.171)	(0.156)
# U.S Months	-	-	-9.99e-05	-8.57e-05
			(0.000165)	(0.000152)
# U.S Trips	-	-	-0.0744	-0.127*
			(0.0733)	(0.0659)
Comm. Effects	-	Yes	Yes	Yes
Migration Effects	-	-	Yes	Yes
Occup. Effects	-	-	-	Yes
Observations	7,566	7,388	7,388	6,329
R-squared	0.000	0.388	0.389	0.509

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

economy. This cannot be said for migrants from Mexico where, if anything, improving English skills is detrimental to labor market outcomes. Evidently, English is more widely spoken in Mexico compared to the rest of L.A because upgrading this skill does not come with any financial benefit – enough people speak English that learning it does not provide a worker with any comparative advantage. Using specification (iv), however, indicates that for no migrant from any country in L.A (that is included in this dataset) do English skills learned abroad significantly contribute to a wage premium at home. Knowing this to be the case, I can conclude that America as a continent is integrated to an extent that makes it possible for a monolinguist to be equally successful as a bilinguist. The fact that the countries within this continent share a profound history might contribute to the ease of immigration; I expect that immigrants from countries in the Middle East or Asia would benefit more from learning English (or whatever language is native in the migration destination) and be at a greater disadvantage if they do not.

VI. Conclusion

Using a testable valid instrument, I determine that a migrant's English skills are improved by staying with a relative when they travel abroad to the U.S, but that this does not have a significant impact on labor market outcomes. Two plausible explanations for the absence of a wage premium for those who are more skilled in speaking English are that English as a skill is not valued in the jobs that Mexicans have at home or that the English learning they engage in while abroad replaces time that could otherwise be spent learning skills that are relevant. The lack of usefulness that speaking English holds could manifest itself in an employer bias. If employers perceive migrants to lack skills that are relevant to their job because they have spent too much time away from home, they may conclude that English speaking has no value and not consider it as a worker differentiator when hiring. If this is true, migrants are at a disadvantage as

workers compared to non-migrants. Even if some individual travels abroad and accumulates human capital (that is not necessarily English skills) that makes them a more valuable worker, employers may not see past the disadvantages of migration.

If an employer bias towards migrants does exist, labor laws should be adjusted to give migrants an advantage in the labor market, or at the least put them on an economically equal plane with non-migrants. Based on the potential for a brain gain, international migration should be encouraged, especially for the less educated, who are the most likely demographic to return to their home country. Somehow encouraging employers to make hiring decisions based on overall relevant job experience, regardless of what country it was gained in, could help to achieve employment equality. Employers should also take note of the usefulness that English learning can have in helping Mexican migrants accumulate human capital. Although improving language skills may not directly cause an increase in wages, they do have the potential to give migrants access to other industry relevant skills – a migrant may be better able to acquire a new skill in their U.S job if they can understand the English explanation that is given to them.

One of the most significant limitations of this study is the measure of English proficiency that I create using the categorical variable provided by the dataset. A binary measure of skill is rather crude and does not capture the complexities of language ability. Future studies of the role that language has in labor market outcomes would ideally have a more robust scale on which to base an individual's ability. Perhaps even a regional-level measure of English would lead to a meaningful estimate, although utilizing such information would depend entirely on the availability of data. The most effective way to understand this phenomenon may be to personally collect the data in a survey focused on language ability. Another notable limitation of my analysis is the IV methodology. Even after demonstrating the validity of who a migrant stays

with as an instrument, it is possible that a different one exists that better explains the labor market outcomes of international migrants. Again, with the appropriate data, other instruments should be tested in future studies. Based on the comparison of results from the MMP and LAMP data sets, I also see the potential to apply my models to countries that do not share the same socioeconomic history that the Americas do. Linguistic difference will be smaller or greater between other countries and its overall relevance in the labor market almost certainly varies among cultures, demonstrating the potential to reveal alternate causal relationships between language and wages.

Appendix

Table 7 – First Stage IV Estimates, English Skills				
Variables	(i)	(ii)	(iii)	(iv)
$\hat{e}$	1.000***	1.001***	0.879***	0.919***
	(0.0308)	(0.0312)	(0.0276)	(0.0247)
Age	-	-1.53e-05	-0.000228	-0.000148
		(0.000788)	(0.000677)	(0.000722)
Age^2	-	-1.42e-07	1.36e-06	6.76e-07
		(7.70e-06)	(6.62e-06)	(7.09e-06)
Male	-	-0.000106	-0.000699	0.000348
		(0.00583)	(0.00501)	(0.00718)
Years of Education	-	-0.000133	0.000350	0.000333
		(0.000559)	(0.000482)	(0.000596)
Job Switch	-	-	0.0201**	0.0134
			(0.00956)	(0.00968)
Relevant Experience	-	-	4.30e-05	-0.00333
			(0.0110)	(0.0117)
# U.S Months	-	-	0.000197**	0.000133*
			(7.85e-05)	(7.64e-05)
# U.S Trips	-	-	-0.000113	0.000334
			(0.00213)	(0.00219)
Comm. Effects	-	Yes	Yes	Yes
Migration Effects	-	-	Yes	Yes
Occup. Effects	-	-	-	Yes
Observations	9,583	7,265	7,265	6,824
R-squared	0.099	0.155	0.377	0.428

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: OLS Estimates of Log Mexican Wages – Migrant, by Occupation					
Variables	(i) *Educator*	(ii) *Supervisor*	(iii) *Skilled Manuf.*	(iv) *Sales*	(v) *Agriculture*
Migrant	0.496	0.00976	0.201	0.917**	0.268
	(0.898)	(0.532)	(0.265)	(0.406)	(0.176)
Male	0.0937	0.924**	0.620***	0.418***	0.176***
	(0.0845)	(0.394)	(0.0633)	(0.0627)	(0.0670)
Years of Education	0.0488***	0.0395***	0.0343***	0.0332***	0.0236***
	(0.00956)	(0.00968)	(0.00604)	(0.00727)	(0.00418)
Age	0.0319	-0.0157	0.0224***	0.0286***	0.00826*
	(0.0229)	(0.0156)	(0.00783)	(0.00934)	(0.00469)
Age^2	-0.000311	0.000172	-0.000247***	-0.000292***	-8.93e-05**
	(0.000245)	(0.000163)	(8.87e-05)	(9.57e-05)	(4.50e-05)
Migrant X Edu	-0.0642**	-0.0321	0.0193	-0.0324	-0.00882
	(0.0292)	(0.0240)	(0.0164)	(0.0223)	(0.0113)
Migrant X Age	0.0232*	0.000885	-0.00474	-0.0187***	-0.00250
	(0.0139)	(0.00863)	(0.00443)	(0.00610)	(0.00265)
Relevant Experience	-	-5.225***	0.0780	0.298	0.0882
		(1.162)	(0.0902)	(0.270)	(0.0636)
# U.S Months	-0.0132***	0.00198	0.000433	4.47e-05	0.00106**
	(0.00485)	(0.00148)	(0.00103)	(0.00104)	(0.000520)
# U.S Trips	0.105	0.0574	-0.0735***	0.00222	-0.0469***
	(0.0828)	(0.114)	(0.0243)	(0.0354)	(0.0145)
Comm. Effects	Yes	Yes	Yes	Yes	Yes
Migrant Effects	Yes	Yes	Yes	Yes	Yes
Observations	313	238	1,198	1,042	2,296
R-squared	0.706	0.658	0.436	0.305	0.506

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: OLS Estimates of Log Mexican Wages – English, by Occupation					
Variables	(i) *Educators*	(ii) *Supervisor*	(iii) *Skilled Manuf.*	(iv) *Sales*	(v) *Agriculture*
English	-0.144	-0.421*	0.121	0.209	0.262***
	(0.346)	(0.213)	(0.121)	(0.156)	(0.0828)
Male	0.0952	0.926**	0.621***	0.421***	0.181***
	(0.0845)	(0.390)	(0.0633)	(0.0628)	(0.0669)
Years of Education	0.0486***	0.0366***	0.0335***	0.0299***	0.0226***
	(0.00955)	(0.00941)	(0.00592)	(0.00713)	(0.00410)
Age	0.0297	-0.0167	0.0221***	0.0293***	0.00818*
	(0.0227)	(0.0153)	(0.00783)	(0.00935)	(0.00465)
Age^2	-0.000290	0.000179	-0.000246***	-0.000310***	-9.19e-05**
	(0.000244)	(0.000161)	(8.87e-05)	(9.57e-05)	(4.49e-05)
Migrant X Edu	-0.0573**	-0.0236	0.0253**	0.00187	-0.00507
	(0.0269)	(0.0164)	(0.0128)	(0.0152)	(0.00881)
Migrant X Age	0.0295***	-0.00221	-0.00198	-0.00643**	0.00105
	(0.0111)	(0.00711)	(0.00287)	(0.00314)	(0.00145)
Relevant Experience		-5.076***	0.0762	0.271	0.116*
		(1.092)	(0.0899)	(0.270)	(0.0635)
# U.S Months	-0.0102*	0.00366**	0.000186	-0.000118	0.000825
	(0.00582)	(0.00168)	(0.00106)	(0.00107)	(0.000523)
# U.S Trips	0.0940	0.0820	-0.0672***	0.0126	-0.0458***
	(0.0901)	(0.111)	(0.0240)	(0.0350)	(0.0142)
Comm. Effects	Yes	Yes	Yes	Yes	Yes
Migrant Effects	Yes	Yes	Yes	Yes	Yes
Observations	313	238	1,198	1,042	2,296
R-squared	0.705	0.665	0.437	0.303	0.508

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10 – IV Estimates of Log Mexican Wages, By Occupation

Variables	(i) *Educators*	(ii) *Supervisor*	(iii) *Skilled Manuf.*	(iv) *Sales*	(v) *Agriculture*
English	-0.0128	0.0964	0.250	0.552	-0.0731
	(1.638)	(0.303)	(0.413)	(0.368)	(0.242)
Age	0.0436**	-0.0258*	0.0187**	0.0271***	0.00803
	(0.0216)	(0.0140)	(0.00811)	(0.0103)	(0.00543)
Age^2	-0.000450*	0.000267*	-0.000205**	-0.000323***	-9.61e-05*
	(0.000232)	(0.000144)	(9.20e-05)	(0.000104)	(5.23e-05)
Male	-0.0456	0.621	0.506***	0.378***	0.206**
	(0.0921)	(0.452)	(0.0705)	(0.0672)	(0.0818)
Years of Education	0.0441***	0.0259***	0.0375***	0.0292***	0.0240***
	(0.0110)	(0.00850)	(0.00610)	(0.00723)	(0.00456)
Job Switch	0.211	-0.348**	-0.103	-0.225*	0.110
	(0.154)	(0.176)	(0.0931)	(0.120)	(0.0715)
# U.S Months	-0.0123	0.00149	0.000123	-0.00133	0.00116**
	(0.0198)	(0.00158)	(0.00140)	(0.00113)	(0.000539)
# U.S Trips	0.147	0.0734	-0.0535**	-0.0113	-0.0622***
	(0.182)	(0.0805)	(0.0258)	(0.0336)	(0.0186)
Comm. Effects	Yes	Yes	Yes	Yes	Yes
Migration Effects	Yes	Yes	Yes	Yes	Yes
Occup. Effects	Yes	Yes	Yes	Yes	Yes
Observations	207	201	934	813	1,667
R-squared	0.668	0.679	0.392	0.295	0.500

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

References

Abarcar, P. (2016). Do employers value return migrants? An experiment on the returns to foreign work experience. *Mathematica Policy Research Reports, 48*

Adams, R., Almeida, H., & Ferreira, D. (2009). Understanding the relationship between founder–CEOs and firm performance. *Journal of Empirical Finance, 16*(1), 136-150.

Adsera, A., & Pytlikova, M. (2015). The role of language in shaping international migration. *Economic Journal, 125*(586), 49.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics, 20*(1), 46-63.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444-455.

Batalova, J., Hallock, J. & Zong, J. (2018). Frequently requested statistics on immigrants and immigration in the United States. Retrieved on February 16, 2018 from https://www.migrationpolicy.org/article/frequently-requested-statistics-immigrants-and-immigration-united-states#Mexican.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review, 77*(4), 531-553.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica, 69*(5), 1127-1160.

Dustmann, C., Fadlon, I., & Weiss, Y. (2011). Return migration, human capital accumulation and the brain drain. *Journal of Development Economics, 95*(1), 58-67.

Greenwood, M. J. & Zahniser, S. S. (1998). Transferability of skills and the economic rewards to US employment for return migrants in Mexico. *Migration between Mexico and the United States. Binational Study, 3*, 1133-1152.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy, 87*(5), 972-990.

Li, S. (2017). Investment and interruption: Effects of the US experience on the earnings of return migrants in Mexico. *Applied Economics*, 1-15.

Lindstrom, D. P. (1996). Economic opportunity in Mexico and return migration from the United States. *Demography, 33*(3), 357-374.

Reinhold, S., & Thom, K. (2013). Migration experience and earnings in the Mexican labor market. *Journal of Human Resources, 48*(3), 768-820.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers, 3*(2), 135-146.

Sadoulet, E. & Winters, P., de Janvry, A. (2001). Family and community networks in Mexico-U.S. migration. *Journal of Human Resources, 36*(1), 159-184.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy, 70*(5), 80-93.

StataCorp. 2015. *Stata 14 Base Reference Manual*. College Station, TX: Stata Press.

Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. *NBER Working Papers Series,* (284)

Wahba, J. (2015). Selection, selection, selection: The impact of return migration. *Journal of Population Economics, 28*(3), 535-563.