



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

The Returns to Wages for Mexican Workers in a Post NAFTA World Has the Gap in Returns to Wages Grown Between Mexican Workers in the U.S. and their Domestic Counterparts?

Ross Hanig

University of Illinois at Urbana-Champaign

Department of Agricultural and Consumer Economics

1301 West Gregory Drive, 405 Mumford Hall, Urbana, IL 61801; rhanig2@uiuc.edu.

Do not cite without author's permission—paper is in working form.

May 2006

Selected Paper prepared for presentation at the American Agricultural Economics
Association Annual Meeting, Long Beach, California, July 23-26, 2006

Abstract

I test to see if the returns to wages for Mexican workers have changed in the Pre- and Post-NAFTA periods using data from the Mexican Migration Project using a pseudo panel analysis framework. The results here suggest that the returns to wages for workers in Mexico have declined significantly in the Post-NAFTA period. However, due to severe measurement error in the data on wages earned in Mexico, and consequently insufficient data to provide suitable cohort sizes, the results here should be viewed as suggestive rather than inferential. I present several tables at the end of the paper with similar regression results from the INEGI data set which do not suffer these problems in measurement error and will form the basis for a significantly updated draft of this paper to be presented at the AAEA conference. The INEGI results suggest the workers in Mexico may actually have not been made significantly worse off from NAFTA. However, the coefficient for workers in the agricultural sector declines, suggesting this sector may have incurred significant losses in wage-returns. These latter results appear consistent with the other work finding that at best, NAFTA may have had an insignificant impact on Mexican wages. The updated version of this paper will present cross-sectional as well as pseudo-panel findings with cohorts grouped by five-year birth spans, gender, and educational attainment to better account for the generational effects and endogeneity with fixed effects.

The Returns to Wages for Mexican Workers in a Post NAFTA World Has the Gap in Returns to Wages Grown Between Mexican Workers in the U.S. and their Domestic Counterparts?

1 Introduction

By most all accounts, labor migration from Mexico to the U.S. has increased since the North American Free Trade Agreement (NAFTA) was enacted in 1994. But in the drive to ratify NAFTA, it seemed policy analysts on both sides of the border thought precisely the opposite would happen. Migrant border crossings to the States were supposed to decline. After all, labor unions expected employers to head south en masse as trade restrictions lifted giving U.S. employers sudden access to the labor surplus that seemed to define the Mexican marketplace. But something unexpected happened. While Mexico indeed appeared to have a substantial advantage to the U.S. in terms of being able to supply low-skill, low-wage labor, it turned out that in an increasingly global marketplace, countries considering changing their trade policies need to look beyond their nearest neighbor to estimate the of those policy changes may have. In Mexico's case, it turned out the country needed to look all the way across the Pacific Ocean to the massive excess labor supply in China. As the Carnegie Endowment for International Peace put it in a policy brief, "[t]o put the size of the global labor oversupply in perspective, if all U.S. jobs were moved to China, there would still be surplus labor in China" (Polaski 2004). So if the U.S. solely traded in cheap goods with Mexico, Mexico could have well expected to reap the gains many predicted. But as the U.S. has ramped up its trading with China, the potential gains could have reasonably been expected come in far smaller than originally forecast.

But learning from the case of Mexico and NAFTA goes far beyond simply learning not to forget about China's potential impact when crafting trade policy. Despite the many important extensive analyses we can turn to in order to learn from NAFTA, (Audley 2004; Beaulieu, Benarroch, and Gaisford 2004; Cragg and Epelbaum 1996;

Deaton and Muellbauer 1980; Gonzalez and Mckinley 1997; Hanson 2003; Polaski 2003, 2004, 2004, 2005; Robertson 2000, 2004), it appears no work to date in the economic literature has empirically explored the effect NAFTA may have had on the returns to wages for Mexicans working both in the United States as well as in Mexico. It is in this area that I contribute to the literature. Additionally, it appears there is a dearth of work accounting for the impact English proficiency may have on wages for Mexican workers both in the U.S. and in Mexico. I account for this potential language impact as well.

Using household level panel data from the Mexican Migration Project (MMP) (Mexican Migration Project (Mmp 107) 2005) I run a series of log-wage regressions with the ultimate goal of attempting to see if on average, a reasonably stable demand for low-wage labor in the U.S. helped draw migrants after NAFTA's enactment when low-price substitute products crowded some of the markets Mexico planned to compete in—causing a less positive-change than expected and possibly even negative. Consequently, since I am interested in focusing on the average worker in this paper, I concentrate on the intercept coefficients as they should help provide estimates of the percent change in wages expected for Mexican workers, controlling for all other variables included. Using a pseudo-panel framework, I first allow each cohort to have its own intercept (testing for joint equality), but restrict the intercepts to be the same both before and after NAFTA's passage. This may provide an indication of whether on average for the period under analysis (1987-2002) Mexican workers received significantly higher wages simply from working in the U.S. and whether these wages significantly differ from one generation to the next. In the second variation, I break each intercept into two-pieces—the before and after NAFTA components. This allows me to compare returns to wages from the country of employment pre- and post within each country. Additionally, it allows me to see if the intercepts by U.S. cohort were roughly equal to their Mexican counterpart in the pre- and post periods. Further, if there does appear to be a significant change in the returns to wages over time by country, this approach allows me to test if the changes were equal—in which case the incentives to migrate would appear to be the same as

in the pre-NAFTA world. However, if the differences differ significantly, this may help explain why migration has changed the way it has, and which workers may have benefited from the policy change.

2 Literature Review

From the previous literature on the Maquiladoras (factories manufacturing intermediate goods in Northern Mexico), it was thought that in NAFTA's wake workers would migrate north where new factories would be established, evening out the wage inequality among Mexico's regions—with wages rising on average (Hanson 1997). Yet this does not appear to have occurred. A number of factors may help explain the discrepancy between theory and reality. First, as has been found in the micro-economic literature, people can form strong social networks to help support themselves in the event they fall on hard times (Munshi and Rosenzweig 2005). Particularly relevant for Mexican labor though, is that with unemployment high, wages low, and rapid change partly spurred by trade liberalization, these likely localized social networks can become very important as informal insurance mechanisms when other forms of feasible insurance are in short supply. Consequently, since migrating within Mexico for potential work can mean losing one's social safety network. This lack of unemployment security may help explain why the rate of convergence between U.S. and Mexico wages actually slowed during the NAFTA period (Robertson 2000). Further, subsidized U.S. agriculture appears to have substantially penetrated the Mexican market contributing to 1.3 million Mexican agricultural workers losing their jobs between 1993 and 2002 (Polaski 2004). And this market penetration was allowed to occur much more quickly than called for in NAFTA.

However, since *ceteris paribus* almost never holds in reality, a number of other changes such the peso crisis and subsequent stabilization may have also significantly

impacted wages. Consequently, the findings here may greatly over-generalize as they fail to account for macro-economic changes. The models here simply account for the pre- and post periods. In future work, it may be especially helpful to account for more macro level changes in developing the testable hypotheses for a model such as those presented here and which estimates may be most likely to suffer from omitted variable bias.

3 Model Development

Where:

$\ln(W_{it})$ = Log Wages for person i in period t

M_{it} = Marital Status

H_i = Human Capital Vector (Educational Attainment)

R_{it} = Job Referral Source (Vector)

P_{it} = Home Community Population (Vector)

A_{it} = Age

E_i = English Proficiency (Vector)

S_{it} = Job Sector (Vector)

G_i = Home Location (Vector)

I begin with the standard log-wage specification:

$$\ln(W_{it}) = \alpha_0 + \beta_A A_{it} + \beta_M M_{it} + \beta_H H_i + \beta_E E_i + \beta_R R_{it} + \beta_S S_{it} + \beta_P P_{it} + \beta_G G_i + \varepsilon_{it} \quad (1)$$

Where the error term is decomposed into its individual and time components leaving a residual:

$$\varepsilon_{it} = \gamma_i + \mu_t + \nu_{it} \quad (2)$$

Equation (1) becomes:

$$\ln(W_{it}) = \alpha_0 + \beta_A A_{it} + \beta_M M_{it} + \beta_H H_i + \beta_E E_i + \beta_R R_{it} + \beta_S S_{it} + \beta_P P_{it} + \beta_G G_i + \gamma_i + \mu_t + \nu_{it} \quad (3)$$

Since leaving α_0 and γ_i as separate terms is likely to yield a final model with potential problems in identifying the parameters of interest, I can add the constant term α_0 to γ_i without having the model suffer. So let:

$$\gamma_{i\alpha} = \alpha_0 + \gamma_i \quad (4)$$

Now adapting the model for pseudo-panel analysis, I sum the terms in equation (3) over all individuals i in each cohort h . This yields:

$$\begin{aligned} \frac{\sum_{i=1}^{I_{h,t}} \ln(W_{ih,t})}{I_{h,t}} = & \frac{\sum_{i=1}^{I_{h,t}} \gamma_{i\alpha h}}{I_{h,t}} + \beta_A \frac{\sum_{i=1}^{I_{h,t}} A_{ih,t}}{I_{h,t}} + \beta_M \frac{\sum_{i=1}^{I_{h,t}} M_{ih,t}}{I_{h,t}} + \beta_H \frac{\sum_{i=1}^{I_{h,t}} H_{ih}}{I_{h,t}} + \beta_E \frac{\sum_{i=1}^{I_{h,t}} E_{ih}}{I_{h,t}} + \beta_R \frac{\sum_{i=1}^{I_{h,t}} R_{ih,t}}{I_{h,t}} \\ & + \beta_S \frac{\sum_{i=1}^{I_{h,t}} S_{ih,t}}{I_{h,t}} + \beta_P \frac{\sum_{i=1}^{I_{h,t}} P_{ih,t}}{I_{h,t}} + \beta_G \frac{\sum_{i=1}^{I_{h,t}} G_{ih}}{I_{h,t}} + \frac{\sum_{i=1}^{I_{h,t}} \gamma_{ih}}{I_{h,t}} + \frac{I_{h,t} \mu_t}{I_{h,t}} + \frac{\sum_{i=1}^{I_{h,t}} \nu_{ih,t}}{I_{h,t}} \end{aligned} \quad (5)$$

In simpler notation, where \tilde{h} denotes the mean value for cohort h at time t , I can express equation (5) as:

$$\begin{aligned} \ln(W_{\tilde{h}t}) = & \gamma_{\tilde{h}\alpha} + \beta_A A_{\tilde{h}t} + \beta_M M_{\tilde{h}t} + \beta_H H_{\tilde{h}} + \beta_E E_{\tilde{h}} + \beta_R R_{\tilde{h}t} + \beta_S S_{\tilde{h}t} \\ & + \beta_P P_{\tilde{h}t} + \beta_G G_{\tilde{h}} + \mu_t + \nu_{\tilde{h}t} \end{aligned} \quad (6)$$

This move from individual data to a pseudo-panel brings in several changes worth noting. First, while having too few observations for proper pseudo-panel analysis has its obvious drawbacks in that my model may be rife with an errors-in-variables problem, allows for far fewer cohorts than is desirable, and contains too few observations to estimate the model I originally intended, there appear to be at least two upsides. The first comes from how I incorporate the ride-hand-side variables that are time-invariant at the person-level. I proxy human capital with a vector of dummy variables corresponding to various levels of educational attainment to allow for the non-linearities in returns to education noted in the micro-development literature (Hanson 2003; Thomas and Strauss 1997). Similarly, I use time-invariant dummies for English proficiency as well as the location of each respondent's home community. And while the marital status dummy can theoretically change over time, in this sample, the dummy is overwhelmingly time-

constant. Consequently estimating a panel of person-specific data with this specification would rule out fixed effects estimation since perfect collinearity would result from least squared dummy variables (LSDV) estimation. And demeaning the data to use a within estimator would yield a parallel effect—this time with the time-invariant variables dropping out of the model entirely.

Meanwhile in a pseudo-panel framework, the variables become proportions for each cohort-year—for example, the proportion of the cohort with secondary education. So variables that previously showed no variation over time for a given panel id, now gain variation since a different group of individuals is included in each new panel id. This appears to be a nice middle-ground in terms of variability for human capital measures between household and nation-level educational attainment data (national attainment proportions exhibit terribly little variation over time since the “cohort” size is roughly equal to that of the nation’s population). But whereas macro data cannot be disaggregated, with pseudo-panel analysis—and given a sufficiently large data set—the researcher can aggregate up. Consequently the researcher can create cohorts with characteristics uniquely of interest yet with slow-moving variables—but likely faster than macro attainment data due to the smaller sample size per cohort—that at the household-level might not move at all¹. Still, while this pseudo-panel analysis appears preferable to a macro approach, it seems implausible that a researcher would aggregate a balanced panel of household data up into a pseudo framework. Primarily, the characteristics of interest that would otherwise be captured at the cohort level could likely be found with a dummy variable—and without having to sacrifice the massive number of observations

¹ While the proportions of education will not vary much within cohort here, as each cohort ages over time, if cohorts were defined at the city or region level for example, we could expect the attainment measures to exhibit more variation over time.

and consequent information required to move into a cohort framework. But with much more household-level data of large cross-sections repeated over time becoming available (Luxembourg Income Study (Lis); McKenzie 2001), this pseudo-panel framework may become increasingly useful for researchers interested in the space where household and macro analysis intersect at the regional level.

An additional change of interest in moving from the household to pseudo-panel framework is that now the fixed effect has an interpretation that may be worth holding onto. Here, the cohort constant corresponds to being male over all cohorts, but further identifies the country of employment, finally separating by five-year spans of birthdates. Consequently, it appears that each fixed effect can be interpreted as the percent change in wages for a male Mexican worker in a particular country simply from being in a particular generational group for the period under analysis (Deaton 1997). The ability this cohort specification presents to glean significantly meaningful interpretation from the intercepts alone is especially helpful here. For example, if we find a positive relationship in one generation's U.S. intercept that is greater than its domestic counterpart, we might expect the next generation's cohort to begin migrating north if they can reasonably observe this pattern.

Since building my cohorts leaves too few observations to run a fully unrestricted model where all coefficients are allowed to vary pre- and post-NAFTA as well as by country, I may be able to simply let the intercepts vary by country, pre- and post-NAFTA and then test for equality. These tests are formulated in the same fashion discussed earlier to see if even after education, job sector, home community and the other characteristics are accounted for, Mexican men in given cohorts could expect higher wages in one

country simply because of its economy. If the differences in intercepts are significant, and if wages are a motivating factor in immigration decisions, this could help shed light on which generations of Mexican workers in general won, lost, or remained the same in terms of expected wages after the change in trade policy.

This slightly less restricted model evolves as follows:

$$\text{Let } \gamma_{\tilde{h}\alpha} = \gamma_{\tilde{h}\alpha}^{\text{Pre}} \text{ if } t < 1994 \quad (7)$$

$$\text{And let } \gamma_{\tilde{h}\alpha} = \gamma_{\tilde{h}\alpha}^{\text{Post}} \text{ if } t > 1994 \quad (8)$$

Then equation (6) becomes:

$$\begin{aligned} \ln(W_{\tilde{h}t}) = & \gamma_{\tilde{h}\alpha}^{\text{Pre}} + \gamma_{\tilde{h}\alpha}^{\text{Post}} + \beta_M M_{\tilde{h}t} + \beta_H H_{\tilde{h}} + \beta_E E_{\tilde{h}} + \beta_R R_{\tilde{h}t} \\ & + \beta_S S_{\tilde{h}t} + \beta_P P_{\tilde{h}t} + \beta_G G_{\tilde{h}} + \mu_t + \nu_{\tilde{h}t} \end{aligned} \quad (9)$$

Where several new hypothesis tests emerge. They are:

- 1) For cohorts in Mexico: $\gamma_{\tilde{h}\alpha}^{\text{Pre}} = \gamma_{\tilde{h}\alpha}^{\text{Post}}$
- 2) For cohorts in the U.S. $\gamma_{\tilde{h}\alpha}^{\text{Pre}} = \gamma_{\tilde{h}\alpha}^{\text{Post}}$
- 3) Pre-NAFTA: $\gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Pre}} = \gamma_{\tilde{h}\alpha, \text{US}}^{\text{Pre}}$
- 4) Post-NAFTA: $\gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Post}} = \gamma_{\tilde{h}\alpha, \text{US}}^{\text{Post}}$

The final question is if the wage-incentive to migrate to and work in the U.S.

grew more than the wage-incentive to work in Mexico. Or in terms of testing:

- 5) If $\gamma_{\tilde{h}\alpha, \text{US}}^{\text{Post}} - \gamma_{\tilde{h}\alpha, \text{US}}^{\text{Pre}} > \gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Post}} - \gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Pre}}$, then migrate.
- 6) If $\gamma_{\tilde{h}\alpha, \text{US}}^{\text{Post}} - \gamma_{\tilde{h}\alpha, \text{US}}^{\text{Pre}} < \gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Post}} - \gamma_{\tilde{h}\alpha, \text{Mexico}}^{\text{Pre}}$, then the overall incentive depends

on the magnitude of change if the period began where $\gamma_{\tilde{h}\alpha, \text{US}}^{\text{Pre}}$ was much greater than

$\gamma_{\tilde{h}\alpha, Mexico}^{Pre}$ or if $\gamma_{\tilde{h}\alpha, US}^{Pre}$ was just slightly greater than $\gamma_{\tilde{h}\alpha, Mexico}^{Pre}$. Still, we would expect this to at the very least provide greater incentive than before to work domestically.

7) If $\gamma_{\tilde{h}\alpha, US}^{Post} - \gamma_{\tilde{h}\alpha, US}^{Pre} = \gamma_{\tilde{h}\alpha, Mexico}^{Post} - \gamma_{\tilde{h}\alpha, Mexico}^{Pre}$, then it appears the wage incentives have not changed since NAFTA's enactment.

While my sample size precludes a fully unrestricted model, I could unrestrict each parameter of interest one at a time, performing the tests above to individually test how the returns to each characteristic may have changed after NAFTA. But since I plan on writing on this topic for my second year paper with data that is sufficient to meet the demands of pseudo-panel estimation, and I hope to develop more specific testable hypotheses from the trade literature, I will leave such estimation for future research.

3.1 Alternative Models (should problems arise in estimation)

Where a standard approach in adapting models like equation (3) for panel analysis is to first-difference the data, bringing in time dynamics and eliminating the fixed effect, that approach poses difficulties given the time invariant data. Further, it would eliminate the fixed effect I very much would like to keep. Additionally, the educational attainment levels that proxy for human capital are time-invariant dummy variables as is the vector for home community location. Still more, since marital status, English proficiency, job referral source, and job sector are dummy variables as well, using a standard first-differencing approach would often take these indicators out of the model entirely—even though they may significantly impact wages in the given time period. As an alternative, I can take an approach similar to that in Bond (2002) and Blundell and Bond (1998) that takes advantage of the possible first degree serial correlation to incorporate time

dynamics. Here, the approach has the additional advantage of allowing the dummy variables to remain in each observation. This can be seen as follows.

Where

$$\nu_{it} = \rho \nu_{i,t-1} + \eta_{it} \quad (10)$$

Then, where η_{it} is assumed to be well behaved, equation (3) becomes

$$\begin{aligned} \ln(W_{it}) = & \alpha_0 + \beta_A A_{it} + \beta_M M_{it} + \beta_H H_i + \beta_E E_{it} \\ & + \beta_R R_{it} + \beta_S S_{it} + \beta_P P_{it} + \beta_G G_i + \gamma_i + \mu_t + \rho \nu_{i,t-1} + \eta_{it} \end{aligned} \quad (11)$$

Noting that the lagged log wage equation is defined as:

$$\begin{aligned} \ln(W_{i,t-1}) = & \gamma_{i\alpha} + \beta_A A_{i,t-1} + \beta_M M_{i,t-1} + \beta_H H_i + \beta_E E_{i,t-1} \\ & + \beta_R R_{i,t-1} + \beta_S S_{i,t-1} + \beta_P P_{i,t-1} + \beta_G G_i + \mu_{t-1} + \nu_{i,t-1} \end{aligned} \quad (12)$$

Then multiplying equation (7) by ρ and subtracting the result from equation (5) yields

the following (where $\rho \nu_{i,t-1}$ cancels out eliminating the serial correlation):

$$\begin{aligned} \ln(W_{it}) = & \rho \ln(W_{i,t-1}) + \beta_A A_{it} - \rho \beta_A A_{i,t-1} + \beta_M M_{it} - \rho \beta_M M_{i,t-1} + \beta_E E_{it} - \rho \beta_E E_{i,t-1} \\ & + \beta_R R_{it} - \rho \beta_R R_{i,t-1} + \beta_S S_{it} - \rho \beta_S S_{i,t-1} + \beta_P P_{it} - \rho \beta_P P_{i,t-1} \\ & + \beta_G G_i (1 - \rho) + \beta_H H_i (1 - \rho) + \mu_t - \rho \mu_{t-1} + \gamma_{i\alpha} (1 - \rho) + \eta_{it} \end{aligned} \quad (13)$$

Now using the following notation in a common factor representation:

$$\begin{aligned} \ln(W_{it}) = & \rho \ln(W_{i,t-1}) + \psi_{A1} A_{it} + \psi_{A2} A_{i,t-1} + \psi_{m1} M_{it} + \psi_{m2} M_{i,t-1} + \psi_{E1} E_{it} + \psi_{E2} E_{i,t-1} \\ & + \psi_{R1} R_{it} + \psi_{R2} R_{i,t-1} + \psi_{S1} S_{it} + \psi_{S2} S_{i,t-1} + \psi_{P1} P_{it} + \psi_{P2} P_{i,t-1} \\ & + \psi_G G_i + \psi_H H_i + \mu_t^* + \gamma_{i\alpha}^* + \eta_{it} \end{aligned} \quad (14)$$

But since I have not yet converted the model into a pseudo-panel framework, the fixed effects would still be perfectly collinear with the time-invariant variables. However, while the time-invariant proxies would be collinear within panel ids, there should be enough variation to prevent them from being terribly collinear in the model as a whole.

However the model is clearly ill-suited to GMM estimation in the mode of Blundell and Bond (1998). Since each time invariant difference would cancel out in the matrix of instruments, these terribly weak instruments could create possibly more bias than might result from the fixed effects being left in the error term. Consequently, if the fixed effect were returned to the error, and the variables were exogenous with respect to the fixed effect but predictably heteroskedastic, the model above could be estimated with random effects and a robust covariance matrix. This is of interest since it appears to be a reasonable way of incorporating time dynamics generally into econometric models with time invariant dummy variables in a fairly straightforward fashion—and without losing the dummies as we would with first differencing.

But in cohorts where the individually time-invariant variables would become slow-moving proportions—yet hopefully fast enough to avoid the collinearity plaguing the normal household panel specification—we have:

$$\begin{aligned} \ln(W_{\tilde{h}t}) = & \rho \ln(W_{\tilde{h},t-1}) + \psi_{A1}A_{\tilde{h}t} + \psi_{A2}A_{\tilde{h},t-1} + \psi_{m1}M_{\tilde{h}t} + \psi_{m2}M_{\tilde{h},t-1} + \psi_{E1}E_{\tilde{h}t} + \psi_{E2}E_{\tilde{h},t-1} \\ & + \psi_{R1}R_{\tilde{h}t} + \psi_{R2}R_{\tilde{h},t-1} + \psi_{S1}S_{\tilde{h}t} + \psi_{S2}S_{\tilde{h},t-1} + \psi_{P1}P_{\tilde{h}t} + \psi_{P2}P_{\tilde{h},t-1} \\ & + \psi_{G1}G_{\tilde{h}t} - \psi_{G2}G_{\tilde{h}t} + \psi_{H1}H_{\tilde{h}t} - \psi_{H2}H_{\tilde{h}t} + \mu_t^* + \gamma_{\tilde{h}\alpha}^* + \eta_{\tilde{h}t} \end{aligned} \quad (15)$$

Where η_{it}^* does not suffer from first order serial correlation, however I have not returned the fixed effect to the error term since the estimate will be of interest when the model is adapted for pseudo-panel analysis (2002). Consequently, if the fixed effects were causing the heteroskedasticiy, there should be none. And if the fixed effects were also the cause of the endogeneity, then placing them in the model should remedy that problem as well. Consequently it appears the model above could be estimated with in a simple fixed effects framework with post-estimation non-linear restrictions tested.

Now letting:

$$\begin{aligned}
\psi_{A1} &= \beta_A & \psi_{m1} &= \beta_M & \psi_{E1} &= \beta_E & \psi_{R1} &= \beta_R \\
\psi_{A2} &= -\rho\beta_A & \psi_{m2} &= -\rho\beta_M & \psi_{E2} &= -\rho\beta_E & \psi_{R2} &= -\rho\beta_R
\end{aligned} \tag{16}$$

$$\begin{aligned}
\psi_{S1} &= \beta_S & \psi_{P1} &= \beta_P & \psi_{G1} &= \beta_G & \psi_{H1} &= \beta_H & \mu_t^* &= \mu_t - \rho\mu_{t-1} \\
\psi_{S2} &= -\rho\beta_S & \psi_{P2} &= -\rho\beta_P & \psi_{G2} &= -\rho\beta_G & \psi_{H2} &= -\rho\beta_H & \gamma_{\tilde{h}\alpha}^* &= \gamma_{\tilde{h}\alpha}(1-\rho)
\end{aligned} \tag{17}$$

Where the following restrictions hold

(and can be extended generally for the right-hand side variables):

$$\begin{aligned}
-\rho\psi_{A1} &= \psi_{A2} & -\rho\psi_{m1} &= \psi_{m2} & -\rho\psi_{E1} &= \psi_{E2} \\
-\rho\psi_{R1} &= \psi_{R2} & -\rho\psi_{P1} &= \psi_{P2}
\end{aligned} \tag{18}$$

We can find the parameters of interest (generally) as follows:

$$\beta_G = \frac{\psi_G}{(1-\rho)} \quad \beta_H = \frac{\psi_H}{(1-\rho)} \tag{19}$$

$$\beta_A = \frac{\psi_{A1} + \psi_{A2}}{(1-\rho)} \Rightarrow \frac{\beta_A + -\rho\beta_A}{(1-\rho)} \Rightarrow \frac{\beta_A(1-\rho)}{(1-\rho)} \tag{20}$$

4 Data

The data from the MMP provides an especially interesting opportunity to uniquely test a number of the hypotheses put forth in the literature that attempt to explain changes in wages for Mexicans since 1994. For example, it is often said that while low-skill Mexican workers have lost out in the post-NAFTA world, medium to high skill workers—especially those in Mexico’s northern region near Maquilladoras—have gained. But with the U.S. and Mexican economies becoming increasingly integrated, it may well be that the returns to English proficiency may be growing and a failure to account for this effect may bias estimates of the effect various levels of education among Mexican workers may have on their wages. And with data on English proficiency, the MMP data allows for these variables to be accounted for in estimation.

However, it appears the MMP dataset, is insufficiently large to estimate a proper pseudo-panel. Even defining my cohorts only by five-year windows of birth, gender, and

country of employment, my cohort sizes still fall short of the 100 observations per cohort year necessary to avoid the errors in estimating sample noted in Verbeek and Nijman (1992). Consequently, the results here should be viewed as exploratory only. Because of data limitations, I have excluded women from the analysis as the cohorts would be based on as few as 2 or 3 observations for most years. In the end, I have just over 2,000 observations with which to make my cohorts. While this is clearly less than ideal for pseudo-panel analysis, these results may yet provide expectations with which to better test in a pseudo-panel for Mexican wages over the time period of interest using the INEGI data set in the next draft of this paper (*Luxembourg Income Study (Lis)*).

Additionally, there appears to be a great deal of error in measurement of wages in Mexico over the years in the survey. This may partly be because of the years of rapid inflation followed by devaluations of the peso, but even in years which were not marked with such high inflation, the mean wages in Mexico differ substantially from the median. And some cases, there appear to be so many extreme values, that even the median wages come in at unbelievably high levels—hundreds or thousands of dollars an hour in wages at times. Since the process I used to convert pesos at current year values to present dollars adjusted for purchasing power parity yielded accurate results when tested with macro data from the World Development Indicators (2005), the problem appears to be with the data. The results are so skewed that even cutting the upper and lower tails still leaves wildly high hourly wages in the data set. Consequently I take the extreme approach of discarding all observations with wages higher than \$100 an hour. And even after taking this extreme approach, the mean wages still differ significantly from the medians as can be seen in Table 9. Consequently, even though my model indicates the mean of the left-hand side variable should be taken for each cohort, I take the median instead in an attempt to yield more accurate results. This may be seen as setting the model up to yield estimates that systematically favor U.S. wages, but I presently do not see a better alternative. Since I remove observations without regard to country though, one way to

look at the results here are as findings for workers who earn less than \$100 an hour. Other variables in the data set do not appear to have these issues.

6 Estimation

6.1 Pre-Estimation Diagnostics

Serial correlation

I fail to reject the null hypothesis of serial correlation at the 1 percent level, but reject the null at the 5 percent level. The data limitations here present an obstacle to accounting for this correlation in the model though. Estimating with the model in equation (15) would require in excess of 60 parameters to be estimated which would cost me about a third of my sample (176 observations) in degrees of freedom. Consequently, any gains in accuracy that might be made in correcting for the serial correlation by manipulating the model could well be expected to be offset by the likely significant bias from the loss in information. As a weak alternative, I use robust covariance estimators where STATA commands allow.

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
F(1, 12) = 5.272
Prob > F = 0.0405

Heteroskedasticity

Not surprisingly, when cohort fixed effects are not accounted for in the model, I reject the null hypothesis of homoskedasticity. However, since I include the fixed effects in the model, this should remedy the problem. Still, the robust estimators should help attenuate any remaining problems of this nature as well.

Likelihood-ratio test

LR chi2(13) = 164.69

(Assumption: . nested in hetero)

Prob > chi2 = 0.0000

Endogeneity With Respect to the Fixed Effects

Performing a random effects regression with the means of the right-hand side variables included, I reject at the one percent level that the variables are exogenous with respect to their individual effects. But as with the problem of heteroskedasticity, including the fixed effects in the model, corrects for this problem as well.

chi2(13)	=	105.37
Prob > chi2	=	0.0000

To help address the problems of estimating with potentially significant errors in the cohort means (using OLS as a comparison set of regressions) I first estimate each specification first using OLS weighted by the number of observations per cohort with a robust covariance matrix that accounts for potential clustering among the error terms by cohort as well. Alternatively, I also use the RREG robust estimator as a median estimator to account for the data quality issues. However, because of the potential serial correlation, I expect the standard errors to be biased downward here, yielding rejections of the null more often than an efficient estimator should. Consequently, I expect the weighted OLS estimator to yield the most plausible results as it best attenuates for what appears to be the most significant problem—uneven cohort sizes and data quality.

7 Results

Even with all the potential problems with data quality and insufficient cohort size, the results here are largely consistent with economic theory and strongly suggest that performing similar tests with the models specified here may yield very interesting results

in a proper pseudo panel framework. From the intercept results reported in tables 1-6, I find that for the entire period under analysis, the expected returns from working in the U.S. were consistently significantly higher for Mexican workers, and that where expected wages appear roughly stable for Mexican workers in the U.S., they appear to have fallen sharply in the Post-NAFTA period. In Table 1, where the intercepts for workers in Mexico were nearly entirely insignificant, the estimates for their counterparts in the U.S. came in positively significant in all models for all cohorts, with the percent increases to wages being the greatest for the youngest cohort. And as seen in Table 2, I reject equality among all cohorts and further consistently reject equality between each cohort and its birth-span counterpart by country.

When I allow the intercepts to differ for the Pre- and Post periods, the intercepts for Mexico that were insignificant when restricted to equality, display a striking pattern when this restriction is not applied. The coefficients for those in Mexico pre-NAFTA come in almost entirely insignificant, save for positively significant estimates in cohorts 3 and 5. But in the Post-NAFTA period all coefficients turn negatively with negative percent income changes in the area of 100 percent for all but one cohort in the weighted regression with similar results from the other estimators. Meanwhile in the U.S. results, the returns to wages from living in the U.S. are insignificant in the Pre- and Post- periods. On the whole, if these results are not biased from the endogeneity that it may be the more entrepreneurial, motivated or highly skilled workers who come to the U.S., this seems to imply a high opportunity cost for Mexican workers who continue to work domestically.

The results reported in Table 5 reinforce the interpretation above in that the U.S. estimates appear insignificantly different before and after NAFTA while the parallel

estimates for Mexico reject equality at the five levels for all except the more middle aged (cohort 4). Further in Table 6, the results indicate that prior to NAFTA the returns to wages simply from geographic location were not significantly different between the U.S. and Mexico. However, after NAFTA's passage, I reject equality at the one percent level in all but one instance, and in that instance I reject equality at the five percent level. And lastly for the intercept results, I reject equality of the difference in returns to wages in the before and after periods at the one percent level for each cohort in Mexico and its match in the U.S. for all but cohort 4, and then I still reject equality in the weighted and OLS regressions. These results suggest that on average, Mexican workers could expect to gain significantly more in terms of avoiding the losses they would suffer in Mexico after NAFTA than before. These results further support Robertson's (2004) finding that wages in the U.S. and Mexico seem to be diverging.

The results for the independent variables reported in Table 7 yield estimates that are largely consistent with expectations. The educational attainment measures repeatedly yield negatively insignificant estimates for primary education which is consistent with prior work (Thomas and Strauss 1997). Then as education levels increase, so too do the coefficients with the college coefficient eventually finding positive significance. The English proficiency measures also come in as expected, with low levels negatively insignificant except a sole instance in the final RREG regression where the sign changes but remains insignificant. It is interesting to note that in the weighted regressions which I expect are the most accurate, the estimates for proficiency increase fairly steadily with each level in both specifications. Meanwhile the increase is somewhat erratic in the OLS and RREG results.

The results for how a job in the U.S. was obtained yields the results most inconsistent with prior work. Along the lines of Munshi (2003) I would expect that obtaining a referral from a close family member or friend would yield positive returns—yet the results here contradict this. The estimates here may be significantly error-ridden though. First, as with the English measures, these are especially likely to be biased as knowing English is likely an endogenous function of motivation as well as a number of other factors not accounted for here. Similarly, Munshi found that a migrants social network in the U.S. which he or she may use to find jobs is a function for many of the rain patters lagged several years (i.e. if a draught is persistent, agricultural workers may migrate, but one bad year may not cause such migration). So measures such as job referral which are related to network size may be significantly biased. And if it is especially low wage workers who use these job referrals, the estimates here may be downwardly biased as they may truly be measuring this lower wage characteristic.

The results for job sector come in largely insignificant with the only discernible pattern being that higher level administrative and supervisory work nets a consistently negative and thrice significant coefficient. However, the estimate only finds significance in the OLS and RREG regressions which do not have robust covariance matrices, so with downwardly biased standard errors, are likely to find significance where none exists. Accordingly, this estimate does not come in significantly in the weighted estimates.

The home community size results may reflect that poverty may be especially present in Mexico's largest cities as these are the only estimates the come in significant, and with extremely large estimates for their percent change in wages—even the smallest

estimate is negative 200 percent change in wages. Lastly, the home community location estimates come in nearly entirely insignificant—save for the first RREG specification.

8 Summary and Conclusion

In this paper, I estimate a series of log wage regressions that strongly suggest future work may be especially fruitful in terms of analyzing the impact NAFTA has had on the returns to wages for Mexican workers. It appears that wages in Mexico appear to be diverging from those in the U.S. when standard free-trade models in the Stolper-Samuelson vein would predict precisely the opposite. In future work it may be helpful to account for models such as that in Davis (1996) that better account for a country's relative factor abundance not simply in general with its closest neighbor, but its relative abundance compared to other countries that produce substitutes for the particular goods a country plans to compete in on the world market. The next draft of this paper (presently in process) will replicate the models tested here concerning wages earned in Mexico—only in the coming version will yield much more reliable results from data on wages in Mexico that do not suffer from the inaccuracies here. As globalization continues to bring the world economy together, research such as this may become increasingly important for policy.

9 References

- Audley, John. 2004. *Nafta's Promise and Reality : Lessons from Mexico for the Hemisphere*. Washington, DC: Carnegie Endowment for International Peace.
- Beaulieu, Eugene, Michael Benarroch, and James Gaisford. 2004. Trade Barriers and Wage Inequality in a North-South Model with Technology-Driven Intra-Industry Trade. *Journal of Development Economics* 75 (1):113-36.
- Blundell, Richard, and Stephen Bond. 1998. Gmm Estimation with Persistent Panel Data: An Application to Production Functions: Institute for Fiscal Studies.
- Bond, Stephen. 2002. Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice: Institute for Fiscal Studies.
- Cragg, Michael Ian, and Mario Epelbaum. 1996. Why Has Wage Dispersion Grown in Mexico? Is It the Incidence of Reforms or the Growing Demand for Skills? *Journal of Development Economics* 51 (1):99-116.
- Davis, Donald R. 1996. Trade Liberalization and Income Distribution: National Bureau of Economic Research, Inc, NBER Working Papers: 5693.
- Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore and London: Johns Hopkins University Press.
- Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge ; New York: Cambridge University Press.
- Gonzalez, Diana Alarcon, and Terry McKinley. 1997. Paradox of Narrowing Wage Differentials and Widening Wage Inequality in Mexico. *Development and Change* 28 (3):505-30.
- Hanson, Gordon H. 1997. Increasing Returns, Trade and the Regional Structure of Wages. *The Economic Journal* 107 (440):113-133.
- , 2003. What Has Happened to Wages in Mexico since Nafta?: National Bureau of Economic Research, Inc, NBER Working Papers: 9563.
- Luxembourg Income Study (Lis)*. Micro database; harmonization of original surveys conducted by the Luxembourg Income Study, Asbl. Luxembourg, periodic updating, 2006. Available from <http://www.lisproject.org/>.
- McKenzie, David J. 2001. Estimation of Ar(1) Models with Unequally Spaced Pseudo-Panels. *Econometrics Journal* 4 (1):89-108.
- Mexican Migration Project (Mmp 107). 2005. Princeton University and the University of Guadalajara.
- Munshi, Kaivan. 2003. Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market. *Quarterly Journal of Economics* 118 (2):549-99.
- Munshi, Kaivan, and Mark Rosenzweig. 2005. Why Is Mobility in India So Low? Social Insurance, Inequality, and Growth?
- Polaski, Sandra. 2003. Trade and Labor Standards: A Strategy for Developing Countries. Washington D.C.: Carnegie Endowment for International Peace.
- , 2004. Carnegie Endowment Policy Brief #30: Job Anxiety Is Real--and It's Global. Washington D.C. : Carnegie Endowment for International Peace.

- , 2004. Mexican Employment, Productivity and Income a Decade after Nafta: Brief Submitted to the Canadian Standing Senate Committee on Foreign Affairs. Washington D.C. : Carnegie Endowment for International Peace.
- , 2005. In Agricultural Trade Talks, First Do No Harm. *Perspectives*:27-30.
- Robertson, Raymond. 2000. Wage Shocks and North American Labor-Market Integration. *American Economic Review* 90 (4):742-64.
- , 2004. Relative Prices and Wage Inequality: Evidence from Mexico. *Journal of International Economics* 64 (2):387-409.
- Thomas, Duncan, and John Strauss. 1997. Health and Wages: Evidence on Men and Women in Urban Brazil. *Journal of Econometrics* 77:159-185.
- Verbeek, Marno, and Theo Nijman. 1992. Can Cohort Data Be Treated as Genuine Panel Data? *Empirical Economics* 17 (1):9-23.
- World Development Indicators, The. 2005. The World Bank Group.

Table 1 Regression Results – Cohort Intercepts

	(1)	(2)	(3)
	OLS	WEIGHTED	RREG
Constant	2.3954 (0.6737)***	1.8402 (0.8255)**	3.3290 (0.5178)***
Working Mexico (Entire Period)			
M1 1949<= Born <1954	0.1882 (0.2913)	0.3381 (0.3323)	0.0153 (0.2239)
M2 1954<= Born <1959	0.3464 (0.2556)	0.4886 (0.2554)*	0.0247 (0.1964)
M3 1959<= Born <1964	0.8023 (0.2656)***	0.8967 (0.2766)***	0.4877 (0.2041)**
M4 1964<= Born <1969	0.3425 (0.2612)	0.4466 (0.2578)	0.0358 (0.2008)
M5 1969<= Born <1974	0.3902 (0.2380)	0.3137 (0.1695)*	0.0139 (0.1829)
M6 1974<= Born <1979	0.2718 (0.2419)	0.2828 (0.1790)	0.1787 (0.1859)
Working U.S. (Entire Period)			
US1 1949<= Born <1954	0.9696 (0.3205)***	1.1713 (0.4382)**	0.8271 (0.2464)***
US2 1954<= Born <1959	0.9717 (0.3127)***	1.0874 (0.4407)**	0.7677 (0.2403)***
US3 1959<= Born <1964	0.9558 (0.3131)***	1.0744 (0.4333)**	0.7322 (0.2406)***
US4 1964<= Born <1969	0.9072 (0.3121)***	1.0567 (0.4144)**	0.6804 (0.2398)***
US5 1969<= Born <1974	0.8882 (0.3013)***	0.9819 (0.3552)**	0.7377 (0.2316)***
US6 1974<= Born <1979	0.9677 (0.3099)***	1.0803 (0.3456)***	0.7255 (0.2382)***
US7 1979<= Born <1984	1.6096 (0.5239)***	1.5668 (0.4731)***	1.4096 (0.4027)***
Observations	176	176	176
R-squared	0.6346	0.6255	0.7271
Standard errors in parentheses			
* significant at 10%; ** significant at 5%; *** significant at 1%			

Table 2 – Cohort Intercepts

	(1)	(2)	(3)
	OLS	WEIGHTED	RREG
All cohort Intercepts Equal	3.21	350.95	5.48
Prob > F	0.0005	0.0000	0.0000
All Mexico Cohort Intercepts Equal	3.71	66.68	5.51
Prob > F	0.0037	0.0000	0.0001
All U.S. Cohort Intercepts Equal	0.51	4.21	1.02
Prob > F	0.8016	0.0143	0.4138
Mexico Cohort 1 = U.S. Cohort 1	17.07	23.71	31.18
Prob > F	0.0001	0.0003	0.0000
Mexico Cohort 2 = U.S. Cohort 2	9.29	5.99	22.21
Prob > F	0.0028	0.0294	0.0000
Mexico Cohort 3 = U.S. Cohort 3	0.56	0.59	2.41
Prob > F	0.4550	0.4579	0.1230
Mexico Cohort 4 = U.S. Cohort 4	8.12	8.50	17.91
Prob > F	0.0052	0.0120	0.0000
Mexico Cohort 5 = U.S. Cohort 5	6.02	10.08	21.52
Prob > F	0.0156	0.0073	0.0000
Mexico Cohort 6 = U.S. Cohort 6	10.22	15.19	10.68
Prob > F	0.0018	0.0018	0.0014

Table 3 – Cohort Intercepts

	(4)	(5)	(6)
	OLS	WEIGHTED	RREG
Constant	2.0827 (0.6468)***	1.7675 (0.7719)**	1.6531 (0.4652)***
Working Mexico (Pre-NAFTA)			
Mex Pre1 1949<= Born <1954	0.2268 (0.3640)	0.1040 (0.2993)	0.1061 (0.2618)
Mex Pre2 1954<= Born <1959	0.2428 (0.3076)	0.2878 (0.1905)	0.0442 (0.2212)
Mex Pre3 1959<= Born <1964	0.7050 (0.3080)**	0.5731 (0.2473)**	0.9490 (0.2215)***
Mex Pre4 1964<= Born <1969	0.1932 (0.2984)	0.1656 (0.2623)	-0.0632 (0.2146)
Mex Pre5 1969<= Born <1974	0.6613 (0.2910)**	0.5531 (0.1622)***	0.6593 (0.2093)***
Mex Pre6 1974<= Born <1979	0.4236 (0.3104)	0.2591 (0.1966)	0.5104 (0.2233)**
Working Mexico (Post-NAFTA)			
Mex Post1 1949<= Born <1954	-0.9614 (0.4567)**	-1.0467 (0.3560)**	-0.7881 (0.3285)**
Mex Post2 1954<= Born <1959	-0.8640 (0.4415)*	-0.9616 (0.3077)***	-0.9689 (0.3176)***
Mex Post3 1959<= Born <1964	-0.6394 (0.4381)	-0.6390 (0.2974)*	-0.5331 (0.3151)*
Mex Post4 1964<= Born <1969	-0.8128 (0.4327)*	-0.9169 (0.2564)***	-0.5603 (0.3112)*
Mex Post5 1969<= Born <1974	-1.0579 (0.4283)**	-1.1642 (0.2557)***	-1.0639 (0.3081)***
Mex Post6 1974<= Born <1979	-0.9593 (0.4229)**	-1.0048 (0.2355)***	-0.9453 (0.3042)***
Mex Post7 1979<= Born <1984	-0.9847 (0.4519)**	-1.0672 (0.3013)***	-0.8747 (0.3250)***

Table 4 – Cohort Intercepts

	(4)	(5)	(6)
	OLS	WEIGHTED	RREG
Working U.S. (Pre-NAFTA)			
US Pre1 1949<= Born <1954	0.2383 (0.2810)	0.1597 (0.2234)	0.2847 (0.2021)
US Pre2 1954<= Born <1959	0.1991 (0.2704)	0.0906 (0.1830)	0.3138 (0.1945)
US Pre3 1959<= Born <1964	0.1035 (0.2665)	0.0163 (0.1644)	0.1513 (0.1917)
US Pre4 1964<= Born <1969	0.1771 (0.2593)	0.0931 (0.1607)	0.2571 (0.1865)
US Pre5 1969<= Born <1974	0.1374 (0.2595)	-0.0096 (0.1343)	0.1711 (0.1867)
Working U.S. (Post-NAFTA)			
US Post1 1949<= Born <1954	-0.0845 (0.4605)	-0.2233 (0.3908)	0.1911 (0.3312)
US Post2 1954<= Born <1959	-0.1909 (0.4433)	-0.3057 (0.3417)	0.0356 (0.3188)
US Post3 1959<= Born <1964	-0.2214 (0.4050)	-0.3371 (0.2683)	-0.0149 (0.2913)
US Post4 1964<= Born <1969	-0.2395 (0.4150)	-0.3425 (0.2735)	-0.0205 (0.2985)
US Post5 1969<= Born <1974	-0.3668 (0.4134)	-0.3764 (0.2793)	-0.1869 (0.2974)
US Post6 1974<= Born <1979	-0.3085 (0.4125)	-0.3555 (0.2766)	-0.1736 (0.2967)
Observations	176	176	176
R-squared	0.7469	0.7357	0.8629

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5 – Cohort Intercepts

	(4)	(5)	(6)
	OLS	WEIGHTED	RREG
All cohort Intercepts Equal	4.39	215.30	12.38
Prob > F	0.0000	0.0000	0.0000
Mex Pre1 = Mex Post1	4.78	5.99	5.24
Prob > F	0.0309	0.0294	0.0240
Mex Pre2 = Mex Post2	5.78	23.28	9.36
Prob > F	0.0179	0.0003	0.0028
Mex Pre3 = Mex Post3	8.28	11.39	19.44
Prob > F	0.0048	0.0050	0.0000
Mex Pre4 = Mex Post4	4.59	10.18	2.16
Prob > F	0.0345	0.0071	0.1441
Mex Pre5 = Mex Post5	14.12	32.44	27.42
Prob > F	0.0003	0.0001	0.0000
Mex Pre6 = Mex Post6	8.58	19.32	18.39
Prob > F	0.0041	0.0007	0.0000
US Pre1 = US Post1	0.48	2.29	0.08
Prob > F	0.4883	0.1543	0.7799
US Pre2 = US Post2	0.68	2.93	0.67
Prob > F	0.4098	0.1109	0.4138
US Pre3 = US Post3	0.53	2.69	0.27
Prob > F	0.4686	0.1247	0.6061
US Pre4 = US Post4	0.86	4.14	0.74
Prob > F	0.3567	0.0628	0.3930
US Pre5 = US Post5	1.21	2.78	1.18
Prob > F	0.2739	0.1195	0.2801

Table 6 – Cohort Intercepts

	(4)	(5)	(6)
	OLS	WEIGHTED	RREG
Mex Pre1 = US Pre1	0.00	0.03	0.59
Prob > F	0.9716	0.8632	0.4429
Mex Pre2 = US Pre2	0.03	2.66	2.39
Prob > F	0.8573	0.1271	0.1253
Mex Pre3 = US Pre3	7.10	8.86	24.15
Prob > F	0.0089	0.0107	0.0000
Mex Pre4 = US Pre4	0.00	0.11	3.37
Prob > F	0.9470	0.7466	0.0690
Mex Pre5 = US Pre5	5.05	16.86	8.47
Prob > F	0.0267	0.0012	0.0044
Mex Post1 = US Post1	19.03	46.66	45.88
Prob > F	0.0000	0.0000	0.0000
Mex Post2 = US Post2	11.12	16.93	47.88
Prob > F	0.0012	0.0012	0.0000
Mex Post3 = US Post3	3.97	10.37	11.79
Prob > F	0.0489	0.0067	0.0008
Mex Post4 = US Post4	8.63	20.51	14.79
Prob > F	0.0040	0.0006	0.0002
Mex Post5 = US Post5	10.72	34.80	33.38
Prob > F	0.0014	0.0001	0.0000
Mex Post1 - Mex Pre1 = US Post1 - US Pre1	5.62	8.26	9.30
Prob > F	0.0195	0.0130	0.0029
Mex Post2 - Mex Pre2 = US Post2 - US Pre2	6.77	70.23	13.75
Prob > F	0.0106	0.0000	0.0003
Mex Post3 - Mex Pre3 = US Post3 - US Pre3	15.37	27.32	49.49
Prob > F	0.0002	0.0002	0.0000
Mex Post4 - Mex Pre4 = US Post4 - US Pre4	4.46	21.63	1.20
Prob > F	0.0370	0.0005	0.2767
Mex Post5 - Mex Pre5 = US Post5 - US Pre5	19.30	150.54	47.10
Prob > F	0.0000	0.0000	0.0000

Table 7 – Independent Variable Results

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	WEIGHTED	RREG	OLS	WEIGHTED	RREG
Education						
Primary	-0.2375 (0.2829)	-0.1074 (0.2733)	-0.4492 (0.2174)**	-0.0600 (0.2651)	-0.0754 (0.2936)	0.1866 (0.1907)
Secondary	0.1757 (0.3091)	0.5054 (0.3693)	0.1947 (0.2376)	0.1850 (0.2822)	0.3069 (0.3581)	0.3999 (0.2030)*
Prep	0.6809 (0.4845)	0.7949 (0.3651)**	0.5222 (0.3724)	0.5575 (0.4622)	0.3574 (0.4594)	0.3005 (0.3325)
College	0.4802 (0.5810)	0.6176 (0.7000)	1.7031 (0.4466)***	0.9880 (0.5856)*	0.9301 (0.4777)*	1.2146 (0.4212)***
English Proficiency						
Do not speak, understand some	-0.2504 (0.3298)	-0.2981 (0.2625)	-0.1815 (0.2535)	-0.0321 (0.2972)	-0.0993 (0.2498)	0.2489 (0.2138)
Do not speak, understand much	0.1682 (0.4591)	0.0921 (0.2920)	-0.1781 (0.3529)	0.5193 (0.4157)	0.3345 (0.2904)	0.8934 (0.2990)***
Speak understand some	0.6097 (0.3800)	0.2597 (0.4555)	0.2918 (0.2920)	0.4399 (0.3487)	0.2327 (0.3637)	0.0593 (0.2508)
Speak understand much	0.0963 (0.6478)	0.3151 (0.6566)	0.2902 (0.4979)	0.8899 (0.6045)	0.6440 (0.6707)	0.9960 (0.4348)**
How Obtained Job in US						
Searched by Oneself	0.1028 (0.3588)	-0.2003 (0.4750)	0.0686 (0.2758)	0.2120 (0.3162)	0.1444 (0.3465)	-0.0899 (0.2274)
Recd by rel, friend, home comm member	-0.7815 (0.2837)***	-0.7631 (0.3725)*	-0.6825 (0.2181)***	-0.0713 (0.2574)	-0.0453 (0.3481)	-0.4218 (0.1851)**
Contracted	0.1137 (0.6200)	-0.0370 (0.7924)	0.4453 (0.4765)	0.0278 (0.5553)	-0.0739 (0.7362)	-0.3594 (0.3994)
Paid friend/home comm member	-1.6152 (6.0768)	-1.2955 (3.2429)	2.5282 (4.6706)	-2.5873 (5.6720)	-0.4039 (2.6126)	-4.1029 (4.0797)
Emp Agency	-0.2220 (1.4292)	-0.6224 (1.3468)	0.0038 (1.0984)	0.7018 (1.2955)	0.0086 (1.6772)	-0.2812 (0.9318)
Street Corner	3.7844 (2.4989)	4.2507 (2.3268)*	2.4339 (1.9207)	2.9558 (2.3143)	3.3147 (1.8467)*	0.5146 (1.6646)

Table 8 Independent Variable Results

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	WEIGHTED	RREG	OLS	WEIGHTED	RREG
Job Sector						
Prof, Tech, Ed, Arts, Perf, Sports	0.9021 (0.6775)	0.9251 (0.8810)	-0.7567 (0.5207)	1.4219 (0.6002)**	1.4313 (0.9019)	1.4403 (0.4317)***
Admin, Super, Direct (Public, Priv)	-2.4169 (1.0495)**	-1.5867 (1.1399)	-2.1868 (0.8066)***	-1.1559 (0.9326)	-0.4367 (1.0852)	-1.6462 (0.6708)**
Ag, Husband, Forest/Fish	-0.4946 (0.5400)	-0.0902 (0.5690)	-0.9433 (0.4151)**	-0.2346 (0.4634)	0.2618 (0.4870)	0.0357 (0.3333)
Manufac/Repair Supers & Equip Ops (Skill)	-0.1248 (0.5626)	0.0955 (0.5780)	-0.2677 (0.4324)	-0.0662 (0.4836)	0.1582 (0.4684)	-0.1395 (0.3479)
Manufac/Repair (Unskill)	-0.5040 (0.5808)	-0.1998 (0.6702)	-0.7666 (0.4464)*	-0.3075 (0.5131)	0.0821 (0.5636)	0.1582 (0.3691)
Admin Support, Sales, Ambulatory, Pers, Direct Serv. etc.	0.1147 (0.5184)	0.2770 (0.5558)	-0.2760 (0.3984)	0.1294 (0.4531)	0.3777 (0.4909)	0.2432 (0.3259)
Home Community Population						
Home Pop 2,500 < 15,000	0.4778 (0.7976)	0.2020 (0.9518)	0.1980 (0.6130)	0.3055 (0.7352)	0.0066 (0.6169)	0.0088 (0.5288)
Home Pop 15,000 < 100,000	0.5600 (0.9719)	0.3670 (1.1213)	0.0633 (0.7470)	0.7243 (0.9197)	0.2536 (0.4796)	1.3670 (0.6615)**
Home Pop 100,000 < 500,000	-5.2709 (1.5803)***	-3.5486 (2.2272)	-3.8378 (1.2146)***	-3.4809 (1.5085)**	-1.9948 (1.5476)	-2.8804 (1.0850)***
Home Community Location						
Mexico Border State	0.2397 (0.5169)	0.1001 (0.6673)	-1.2442 (0.3973)***	0.5312 (0.4728)	0.3180 (0.7968)	0.5287 (0.3401)
Mexico Northern State	-0.5375 (0.4108)	-0.5662 (0.3869)	-1.4550 (0.3157)***	-0.3003 (0.3822)	-0.3276 (0.4439)	0.1005 (0.2749)
Mexico Middle State	-0.2249 (0.4054)	-0.2771 (0.4534)	-0.8792 (0.3116)***	-0.1860 (0.3876)	-0.1047 (0.3782)	-0.1486 (0.2788)
Mexico Southern State	0.8445 (0.5874)	0.7010 (0.9322)	-0.5927 (0.4515)	0.3199 (0.5491)	0.2768 (0.9234)	-0.0392 (0.3949)

Table 9
(All values are in means unless otherwise noted)

	Mexico (Place of Work)					
	M1	M2	M3	M4	M5	M6
	1949<= Born <1954	1954<= Born <1959	1959<= Born <1964	1964<= Born <1969	1969<= Born <1974	1979<= Born <1984
Year Born	1951.169	1956.17	1961.378	1965.956	1971.011	1975.61
Median Wage Plausible	\$3.61	\$3.63	\$7.43	\$4.92	\$5.17	\$4.31
Mean Wage Plausible	\$16.05	\$12.52	\$18.29	\$12.89	\$14.33	\$14.40
age	46.8814	40.6863	35.4508	30.8496	26.0055	21.7524
married	0.8729	0.8235	0.7824	0.7035	0.6066	0.4286
Primary	0.2458	0.2876	0.2694	0.3186	0.1967	0.2571
Secondary	0.0508	0.0850	0.1503	0.1770	0.3388	0.2952
Prep	0.0424	0.0392	0.0674	0.1018	0.1038	0.0762
College	0.0508	0.0850	0.1140	0.0752	0.0765	0.0476
Do not speak, understand some	0.2119	0.1699	0.2798	0.1858	0.1475	0.1714
Do not speak, understand much	0.0763	0.0719	0.0829	0.1150	0.0601	0.1143
Speak understand some	0.1102	0.1046	0.0829	0.0841	0.1311	0.0667
Speak understand much	0.0085	0.0131	0.0155	0.0133	0.0109	0.0381
Searched by Oneself	0.2034	0.1373	0.1554	0.1549	0.0984	0.1810
Recd by rel, friend, home comm member	0.3814	0.3660	0.4093	0.3584	0.3388	0.2857
Contracted	0.0508	0.0392	0.0363	0.0265	0.0492	0.0190
Paid friend/home comm member	0.0000	0.0000	0.0000	0.0088	0.0000	0.0000
Emp Agency	0.0000	0.0000	0.0000	0.0044	0.0055	0.0095
Street Corner	0.0085	0.0000	0.0000	0.0000	0.0000	0.0000
Prof, Tech, Ed, Arts, Perf, Sports	0.0678	0.0980	0.1295	0.0841	0.0710	0.0286
Admin, Super, Direct (Public, Priv)	0.0169	0.0327	0.0259	0.0177	0.0219	0.0190
Ag, Husband, Forest/Fish	0.3051	0.2288	0.1762	0.1195	0.1858	0.1429
Manufac/Repair Supers & Equip Ops (Skill Work)	0.2458	0.2157	0.2953	0.3274	0.3607	0.3333
Manufac/Repair (Unskill)	0.1017	0.1503	0.0829	0.1416	0.0874	0.1714
Admin Support, Sales, Ambulatory, Pers, Direct Serv. etc.	0.2034	0.1699	0.2073	0.1991	0.2022	0.2381
Home Pop 2,500 < 15,000	0.0678	0.0327	0.0881	0.0752	0.0601	0.0476
Home Pop 15,000 < 100,000	0.0000	0.0000	0.0052	0.0177	0.0109	0.0190
Home Pop 100,000 < 500,000	0.0085	0.0000	0.0104	0.0044	0.0000	0.0000
Home Pop 100,000 < 500,000	0.8729	0.9412	0.8705	0.8761	0.8852	0.9048
Mexico Border State	0.1695	0.1830	0.2073	0.1814	0.2678	0.2952
Mexico Northern State	0.2542	0.2876	0.2746	0.2257	0.1967	0.1524
Mexico Middle State	0.4746	0.4183	0.3834	0.4646	0.4536	0.4095
Mexico Southern State	0.0169	0.0261	0.0104	0.0133	0.0219	0.0095

Table 10 (Values in means unless noted otherwise)	US (Place of Work)						
	US1 1949<= Born <1954	US2 1954<= Born <1959	US3 1959<= Born <1964	US4 1964<= Born <1969	US5 1969<= Born <1974	US6 1979<= Born <1984	US7 1984<= Born<1989
Year Born	1951.134	1956.031	1961.195	1966.048	1970.899	1975.521	1980.235
Median Wage Plausible	\$6.51	\$7.01	\$6.85	\$7.15	\$6.62	\$7.03	\$7.89
Mean Wage Plausible	\$8.21	\$8.75	\$8.48	\$8.71	\$7.85	\$8.57	\$10.17
age	41.0369	35.6849	30.8947	26.3287	22.5629	20.5966	18.6471
married	0.9263	0.8938	0.8316	0.7195	0.5000	0.3866	0.3529
Primary	0.3318	0.3185	0.3447	0.2759	0.2956	0.3697	0.4118
Secondary	0.0599	0.0993	0.1500	0.2046	0.2893	0.2437	0.2941
Prep	0.0184	0.0342	0.0474	0.0966	0.0881	0.0924	0.1765
College	0.0046	0.0411	0.0684	0.0322	0.0031	0.0000	0.0000
Do not speak, understand some	0.2811	0.2979	0.2789	0.2667	0.3365	0.2941	0.1765
Do not speak, understand much	0.0922	0.0685	0.1158	0.1425	0.1384	0.1849	0.1765
Speak understand some	0.2765	0.2568	0.2684	0.2920	0.1761	0.2605	0.1765
Speak understand much	0.0599	0.1130	0.0947	0.0621	0.0755	0.0672	0.2353
Searched by Oneself	0.2903	0.2705	0.2368	0.2690	0.2327	0.1681	0.2941
Recd by rel, friend, home comm member	0.5945	0.6370	0.7053	0.6460	0.7044	0.7227	0.7059
Contracted	0.0553	0.0479	0.0263	0.0368	0.0346	0.0756	0.0000
Paid friend/home comm member	0.0000	0.0000	0.0000	0.0023	0.0000	0.0000	0.0000
Emp Agency	0.0046	0.0000	0.0000	0.0115	0.0031	0.0084	0.0000
Street Corner	0.0000	0.0068	0.0000	0.0023	0.0000	0.0000	0.0000
Prof, Tech, Ed, Arts, Perf, Sports	0.0000	0.0103	0.0053	0.0069	0.0031	0.0168	0.0000
Admin, Super, Direct (Public, Priv)	0.0046	0.0034	0.0000	0.0023	0.0000	0.0084	0.0000
Ag, Husband, Forest/Fish	0.3548	0.2979	0.2316	0.2345	0.1981	0.1681	0.0588
Manufac/Repair Supers & Equip Ops (Skill Work)	0.1659	0.2123	0.2184	0.2506	0.2296	0.3361	0.4706
Manufac/Repair (Unskill)	0.2212	0.2432	0.2711	0.2598	0.2484	0.2269	0.2941
Admin Support, Sales, Ambulatory, Pers, Direct Serv. etc.	0.2304	0.2158	0.2526	0.2230	0.3082	0.2269	0.1765
Home Pop 2,500 < 15,000	0.0230	0.0548	0.0342	0.0345	0.0377	0.0336	0.0588
Home Pop 15,000 < 100,000	0.0369	0.0479	0.0342	0.0368	0.0314	0.0252	0.0588
Home Pop 100,000 < 500,000	0.0000	0.0068	0.0105	0.0161	0.0094	0.0000	0.0000
Home Pop 100,000 < 500,000	0.9032	0.8493	0.8684	0.8644	0.8679	0.9244	0.8824
Mexico Border State	0.0276	0.0548	0.0789	0.0874	0.1352	0.2269	0.0000
Mexico Northern State	0.3687	0.2979	0.3105	0.2690	0.2358	0.2689	0.2941
Mexico Middle State	0.5392	0.5753	0.5395	0.5816	0.5094	0.4454	0.6471
Mexico Southern State	0.0507	0.0582	0.0500	0.0230	0.0755	0.0000	0.0000

Table 11
INEGI Data
Log(Wage) in Pesos is dependent variable

	Coef.	Robust Std. Err.	P>t
ln(age)	0.1535	0.0264	0.0000
Sex	0.2590	0.0139	0.0000
Married	0.0516	0.0130	0.0000
Primary	0.2973	0.0156	0.0000
Secondary	0.4807	0.0191	0.0000
Prepatory	0.7630	0.0218	0.0000
Superior	1.0914	0.0303	0.0000
Profession	0.6152	0.0395	0.0000
Tech	0.5367	0.0309	0.0000
Education	0.8281	0.0301	0.0000
Art, show busines and sports	0.6809	0.0558	0.0000
Supervisors & Directors	0.8860	0.0297	0.0000
Ag	-0.5642	0.0262	0.0000
Heads and supervisors in crafts industry	0.5405	0.0388	0.0000
Craftsmen, drivers and mobile plant operators	0.1574	0.0217	0.0000
Operators (assembly line),labourers in craft or manufacturing	0.0330	0.0232	0.1550
Admin	0.4533	0.0252	0.0000
Shopkeepers, shop employees and sales agents	0.1233	0.0283	0.0000
Personal and Domestic Services	0.0603	0.0257	0.0190
Border	0.2079	0.0130	0.0000
Noth	-0.0642	0.0176	0.0000
Capital	0.1283	0.0159	0.0000
South	-0.2354	0.0152	0.0000
Urban	0.3112	0.0176	0.0000
1989	0.0028	0.0273	0.9190
1992	-0.0079	0.0276	0.7740
1994	0.0428	0.0268	0.1100
1996	-0.3601	0.0272	0.0000
1998	-0.3156	0.0273	0.0000
2000	-0.2207	0.0281	0.0000
2002	-0.1653	0.0267	0.0000
Constant	1.0120	0.1166	0.0000

Table 12
INEGI Data
Log(Wage) in Pesos is dependent variable
Self- Employed Removed

		Robust		
Self- Employed Removed		Coef.	Std. Err.	P>t
ln(age)		0.1626	0.0179	0.0000
Sex		0.2514	0.0133	0.0000
Married		0.0534	0.0118	0.0000
Primary		0.2882	0.0130	0.0000
Secondary		0.4647	0.0160	0.0000
Preparatory		0.7532	0.0195	0.0000
Superior		1.0866	0.0289	0.0000
Profession		0.6619	0.0396	0.0000
Tech		0.5596	0.0294	0.0000
Education		0.8219	0.0300	0.0000
Art, show busines and sports		0.7169	0.0559	0.0000
Supervisors & Directors		0.9023	0.0296	0.0000
Ag		-0.4374	0.0252	0.0000
Heads and supervisors in crafts industry		0.5497	0.0385	0.0000
Craftsment, drivers and mobile plant operators		0.1878	0.0212	0.0000
Operators (assembly line),labourers in craft or manufacturing		0.0321	0.0231	0.1640
Admin		0.4549	0.0248	0.0000
Shopkeepers, shop employees and sales agents		0.1967	0.0280	0.0000
Personal and Domestic Services		0.0870	0.0246	0.0000
Border		0.2122	0.0122	0.0000
Noth		-0.0360	0.0163	0.0280
Capital		0.1311	0.0127	0.0000
South		-0.2085	0.0148	0.0000
Urban		0.2716	0.0158	0.0000
	1989	-0.0480	0.0222	0.0310
	1992	-0.0506	0.0223	0.0230
	1994	-0.0096	0.0215	0.6540
	1996	-0.4052	0.0209	0.0000
	1998	-0.3568	0.0215	0.0000
	200	-0.2598	0.0232	0.0000
	2002	-0.2144	0.0215	0.0000
Constant		1.0650	0.0742	0.0000

Table 13
Pre-NAFTA

	Coef.	Robust Std. Err.	P>t
ln(age)	0.1646	0.0692	0.0170
Sex	0.2567	0.0270	0.0000
Married	0.0716	0.0389	0.0650
Primary	0.3064	0.0342	0.0000
Secondary	0.4846	0.0451	0.0000
Preparatory	0.7160	0.0458	0.0000
Superior	0.9581	0.0498	0.0000
Profession	0.5838	0.0575	0.0000
Tech	0.4608	0.0540	0.0000
Education	0.7359	0.0512	0.0000
Art, show busines and sports	0.6129	0.0843	0.0000
Supervisors & Directors	0.9219	0.0508	0.0000
Ag	-0.5857	0.0488	0.0000
Heads and supervisors in crafts industry	0.5561	0.0589	0.0000
Craftsment, drivers and mobile plant operators	0.1876	0.0376	0.0000
Operators (assembly line),labourers in craft or manufacturing	0.0446	0.0469	0.3420
Admin	0.4886	0.0441	0.0000
Shopkeepers, shop employees and sales agents	0.1741	0.0520	0.0010
Personal and Domestic Services	0.0091	0.0533	0.8640
Border	0.1591	0.0246	0.0000
Noth	-0.1214	0.0357	0.0010
Capital	0.0599	0.0373	0.1080
South	-0.1472	0.0293	0.0000
Urban	0.3222	0.0298	0.0000
1989	0.0016	0.0252	0.9510
1992	-0.0164	0.0254	0.5180

Table 13
Post-NAFTA

Post-NAFTA	Robust			
	Coef.	Std. Err.	P>t	
ln(age)	0.1510	0.0258	0.0000	
Sex	0.2603	0.0177	0.0000	
Married	0.0425	0.0142	0.0030	
Primary	0.2932	0.0182	0.0000	
Secondary	0.4845	0.0210	0.0000	
Preparatory	0.7755	0.0263	0.0000	
Superior	1.1362	0.0410	0.0000	
Profession	0.5931	0.0553	0.0000	
Tech	0.5473	0.0419	0.0000	
Education	0.8455	0.0407	0.0000	
Art, show busines and sports	0.6856	0.0843	0.0000	
Supervisors & Directors	0.8482	0.0403	0.0000	
Ag	-0.5553	0.0337	0.0000	
Heads and supervisors in crafts industry	0.5334	0.0553	0.0000	
Craftsment, drivers and mobile plant operators	0.1201	0.0297	0.0000	
Operators (assembly line),labourers in craft or manufacturing	0.0290	0.0305	0.3420	
Admin	0.4187	0.0335	0.0000	
Shopkeepers, shop employees and sales agents	0.0853	0.0376	0.0230	
Personal and Domestic Services	0.0562	0.0326	0.0840	
Border	0.2521	0.0172	0.0000	
Noth	-0.0284	0.0199	0.1530	
Capital	0.1612	0.0174	0.0000	
South	-0.2830	0.0193	0.0000	
Urban	0.3171	0.0256	0.0000	
	1996	-0.3213	0.3123	0.3040
	1998	-0.2770	0.3127	0.3760
	2000	-0.1788	0.3127	0.5670
	2002	-0.1236	0.3133	0.6930