Evaluating the Effects of Legalization on Farmworker Wages in the Crop Sector

Chellie A. Hogan  
Department of Agricultural Economics  
Mississippi State University  
Ch2780@msstate.edu

Lurleen M. Walters  
Department of Agricultural Economics  
Mississippi State University  
Lurleen.Walters@msstate.edu

Brian Williams  
Nebraska Public Power District  
brwilli@nppd.com

Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-7, 2018

Copyright 2018 by Chellie A. Hogan, Lurleen M. Walters, and Brian Williams. All rights reserved.  
Readers may make verbatim copies of this document for non-commercial purposes by any means,  
provided that this copyright notice appears on all such copies.
Abstract

Labor intensive sectors such as the specialty crop sector have historically had strong reliance on foreign labor. It is estimated that more than half unauthorized of the foreign-born labor force in the specialty crop sector are unauthorized for US employment, and it is often argued that this situation is untenable and risky for producers. However, proposals suggesting authorization for current unauthorized farm workers tends to generate controversy on the grounds that it may signal a reward for illegal entry and employment, and violation of federal immigration laws. Using data from the National Agricultural Workers Survey for 1989-2014, this study utilizes a propensity score matching and minimum-biased treatment effects estimation methods to evaluate the farm wage implications of authorization of foreign-born specialty crop farm workers. We compare findings for the national level versus California. Positive wage effects are estimated nationally and in California, with higher magnitude effects observed in California.
Introduction

Foreign born workers have consistently comprised a significant proportion of the hired labor force in the U.S. agricultural industry over the past several decades. This has been most evident in the dairy, meatpacking and specialty crop sectors (Martin, 2017; Artz, Orazem, and Otto, 2007; Krumel, 2017). The specialty crop industry, specifically the fruit, vegetable and horticultural (FVH) sector, is unique in that it requires the most labor relative to all other agricultural inputs. In this sector, labor availability impacts the choice of production technologies, cropping patterns, and the competitiveness of U.S. producers relative to low-cost foreign producers (Boucher and Taylor, 2007).

According to the 2012 Census of Agriculture, FVH farms constituted 21% of total U.S. farms, however, they accounted for 63% of the $19 billion of direct-hire expenses and 80% of the $5.3 billion of contract crop labor expenses (Martin, 2014). FVH farms had the highest hired labor costs in 2016, with vegetable farms, fruit and nut farms, and greenhouse/nursery farms reporting an estimated $3.2 billion, $5.5 billion, and $4.7 billion, respectively, in hired labor costs (Rural Migration News, 2017). Other crops such as soybeans, cotton, and corn, have been able to successfully implement mechanization and have experienced increases in productivity and decreases in labor costs. In comparison, FVH farms tend to use labor intensively and have had fewer options for mechanization given the nature of their products, which are usually ill suited for such advancements. Many specialty crops, particularly fresh fruits and vegetables, are sold to a consumer base that demands produce with a minimal damage, bruising, and blemishes. Further, these crops tend to be highly perishable and have short windows during which produce must be harvested. Mechanization adoption in the U.S. FVH subsector has hinged upon benefit-cost decisions in which the cost and availability of farm labor is taken into account.
Implementing mechanization requires large capital commitments, changes in field configurations, and potentially an overall change in the farming operation (Huffman, 2012).

These aforementioned factors appear to have disincentivized innovation in suitable mechanized options and encourage reliance on farm labor. Much of the farm labor force lacks authorization for US employment. The risk that this poses is not lost on producers, and it has sparked much concern among producers/producer groups to date. According to Krogstad, Passel, and Cohn (2017), undocumented immigrants are overrepresented in farming, constituting 26% of all farm workers, when compared to their estimated 5% share of the civilian labor force. In terms of hired labor on crop farms however, the representation is more pronounced: according to the 2013-2014 National Agricultural Workers Survey (NAWS) findings, 73 percent of all hired crop farm workers were foreign born, and only 53 percent of these workers were authorized to work in the United States (Hernandez, Gabbard, and Carroll, 2016).

With the changes in immigration trends and the current political environment, farmers are understandably concerned that immigration reform, especially regarding the increased stringency of immigration enforcement may disrupt labor supply and curb their access to workers. Fewer workers could lead to an increase in labor costs and adversely affect U.S. producers’ competitiveness in the global market for FVH products (Calvin and Martin, 2010). Given that substantial proportions of the workforce are foreign born and unauthorized for US employment, it is important to consider the effects that changes in immigration policy may have.

This study uses a treatment effects framework to evaluate the farm wage implications of authorization, and to make comparisons between the national and regional level earnings outcomes. This study is similar to previous work (Walters, Emerson, and Iwai, 2008; Kandilov and Kandilov, 2010) that have argued that farm workers self-select into legal status, but goes
further to assess the outcomes on national and regional levels. Specifically, we evaluate the impact of legal status on farm worker earnings outcomes for authorized and unauthorized workers in the U.S. crop farm workforce, and assess the differences in potential earnings outcomes at the national and regional levels via propensity score matching and bias minimizing treatment effects estimation techniques. This study uses data for 1989-2014 from the National Agricultural Workers Survey (NAWS), which is a rich nationally representative data set. It is an employment-based, random-sample survey of U.S. crop workers that collects demographic, employment, and health data in face-to-face interviews. It also collects data on the legal status of survey respondents.

Following this introduction, section two reviews the relevant literature. Section three describes the methods and model specification, and section four discusses the results. The study concludes with a discussion of the findings and potential future research directions.

**Literature Review**

Studies that have analyzed wage effects of authorization have determined that obtaining legal status has a positive effect on the average earnings of foreign-born workers. They found that unauthorized immigrants experience an estimated 5 to 15 percent wage penalty when compared to their authorized counterparts (Taylor, 1992; Cobb-Clark, Shiells, and Lowell, 1995; Rivera-Batiz, 1999; Kossoudji and Cobb-Clark, 2002; Iwai, Emerson and Walters, 2006; Walters, Emerson and Iwai, 2008; Kandilov and Kandilov, 2010; Nisbet and Rodgers III, 2013; Borjas, 2017). For brevity, we focus on the findings of studies that focused specifically on foreign-born farmworkers.
Taylor (1992) found that unauthorized immigrants select into low-skill, low-paying farm jobs and therefore experience lower earnings than their authorized counterparts (using data from a 1983 survey conducted by the University of California and the California Employment Development). Taylor explains, based on the theory of human capital developed by Chiswick (1978) that the lack of legal status creates a barrier to higher earnings and increased job mobility due to the unauthorized immigrant worker prioritizing the minimization of detection rather than maximization of earning potential. Taylor employed a two-step procedure proposed by Heckman (1979), found that unauthorized workers earn less than authorized workers, and concluded that that lack of legal status serves as a barrier to job mobility, and adversely affects earnings potential.

In their study, Isé and Perloff (1995) separated legal status into five different categories and used a multinomial logit to explain the probability of a certain legal status as a function of demographic characteristics. Their analysis utilized a subset of the National Agricultural Workers Survey data (1989-1991) relied on Lee’s extension of Heckman’s two-step procedure to mitigate selection bias and obtain consistent wage estimates. The authors found that authorized foreign-born workers earn 15 percent more, on average, per hour and per week than unauthorized workers (Isé and Perloff, 1995).

Emerson (2007) discussed how wage rates could be affected under alternative immigration policy scenarios. He argued that if it were the case that all workers were authorized for US employment via full authorization and producers had access to guest workers, but that no changes in technology occurred (so that the structure of labor demand remained unchanged), then market-determined wage rates could be expected to remain at the levels observed for the legal workers. The only difference would be the absence of a wage penalty for the formerly
undocumented workers and higher direct wage costs for employers (Emerson, 2007).

Alternatively, if all undocumented workers were removed with limited or no replacement guest workers and borders were closed, there could be increased agricultural wages given fixed technology and product mix, and capital immobility, at least in the short term (Emerson, 2007). Alluding to the benefits of temporary or permanent immigrant labor that could boost the U.S. economy, Emerson argued that immigrant workers and complementary factors of production, including capital, land and complementary labor, would capture the gains, while losses would be absorbed by substitute labor.

Iwai, Emerson, and Walters (2006) used data from the National Agricultural Workers Survey for 1989-2004 to estimate U.S. farmworker wage differentials by legal status via a Heckman-type two-stage estimation method to control for selection bias issues. They determined that workers that were authorized for employment in US farm work had earned higher wages on average than those who lacked authorization for US employment.

Walters, Emerson, and Iwai (2008) used a treatment effects approach that utilized parametric and nonparametric methods to evaluate the wage implications of legal status for foreign born crop workers in the United States. In this context, legal status was used as the “treatment” to assess the potential earnings outcomes for authorized versus unauthorized foreign born farmworkers. The study used National Agricultural Workers Survey data for 1989 to 2006. The study findings indicated a positive wage effect due to legal status that was consistent with findings by previous research (Taylor, 1992; Isé and Perloff, 1995; Iwai, Emerson, and Walters, 2006).

Kandilov and Kandilov (2010) evaluated the effect of legal status on wages and health insurance of foreign-born farmworkers utilizing data for 2000-2006 from the National
Agricultural Workers Survey. The authors employed propensity score matching techniques and focused on a subset of single male farmworkers that were employed fulltime in US farm work. They reported a wage gain of roughly 5 percent in average wages of undocumented workers due to authorization, which is consistent with previous studies’ findings.

Nisbet and Rodgers III (2013) argued that the tweaks to U.S. immigration policy, social policy, and overall migration trends since the implementation of IRCA have shaped the farm labor market, specifically the structure of wages, for both authorized and unauthorized U.S. farm workers. Using National Agricultural Workers Survey (NAWS) data spanning from 1990-2009, the authors estimated a log wage function to estimate the unadjusted wage gap between the authorized and unauthorized farmworkers. The authors then identified the key predictors of the wage differential between authorized and unauthorized workers, by utilizing decomposition techniques\(^1\). Finally, the authors decomposed the changes in the wage gap across time. They concluded that IRCA changed the nature of the labor market in the agricultural sector—wage differences grew and returns to measured characteristics differed\(^2\) according to legal status (Nisbet and Rodgers III, 2013).

While the previous studies have been focused at the national level, this study has an additional component that looks at the potential effects at the regional level, and it assesses these over a longer time frame. In terms of the overall methodological approach, it is similar to work by Walters, Emerson and Iwai (2008) and Kandilov and Kandilov (2010). However, whereas

\(^1\) Used in Oaxaca (1973)
\(^2\) The authors used the Juhn Murphy Pierce (1991) decomposition. Measured characteristics included: years of schooling, job tenure, English speaking ability, age, reading ability, migration status, and demographic measures such as birthplace, race, gender, marital status, and number of children. Significant characteristics included: job tenure, English speaking ability, years of schooling and age.
Walters, Emerson, and Iwai (2008) utilized the full data for 1989 to 2006 and focused on earnings outcomes, Kandilov and Kandilov (2010) restricted their analysis to a six year sample of unmarried male farmworkers and focused on the effects on earnings and benefits. This study uses propensity score matching (PSM) and minimum-biased treatment effects (MBTE) estimation methods. The latter was proposed by Millimet and Tchernis (2013) and is used as a robustness check of the PSM. The MBTE method does not appear to have been previously applied in the context of farm labor and immigration policy research.

While the previous studies have been focused at the national level, our study has an additional component that looks at the potential effects at the regional level, and it assesses these over a longer time frame. In terms of the overall methodological approach and assessing the treatment effects, it is similar to work by Walters, Emerson and Iwai (2008) and Kandilov and Kandilov (2010). However, Kandilov and Kandilov (2010) restricted their analysis to a six year sample of unmarried male farmworkers, and only used propensity score matching to analyze the wage effects of legalization. Our study does not have this restriction.

Finally, while our study uses propensity score matching (PSM) as used in previous work, it also uses a minimum-biased estimation (MBE) method proposed by Millimet and Tchernis (2013) to check robustness of the PSM. Whereas PSM is appropriate for addressing selection bias from observed factors, the MBE is appropriate where selection bias arises from factors that are unobservable to researchers. In the context of this research, education and experience are examples of characteristics (factors) that would typically be observed. Characteristics that are unlikely to be observed are immigrant worker’s motivations for migrating to, working in the United States and/or gaining legal status (or conceivably the factors underlying such motivations), how immigrant workers may perceive being affected by changes in US
immigration policy, or how they may modify their behaviors in response to media reports on immigration policy changes. The MBE method was not used in either of the previous studies, and to the best of my knowledge, it has not been applied in the context of farm labor and immigration policy research.

**Methods**

**Treatment Effects Framework**

The treatment in this study is defined as the authorization of unauthorized workers in the specialty crop sector. Because the treatment variable (legal status) is binary, making it impossible to observe the same individual in both states at the same time; causal connections must be inferred using counterfactuals. In the context of this study, the wage effect (outcome) of a change in legal status (treatment) is of primary interest. Since we are interested mainly in the wage effect caused by the ‘switching’ of the binary treatment variable, causation is identified *ceteris paribus*, signifying that all other independent variables are held constant. There is a participant group \((D^* = 1)\) (treatment) (authorized) and a non-participant group \((D^* = 0)\) (untreated) (unauthorized), with the treatment effect being denoted as the causal effect of the treatment on the outcome (Long, 1997). The value in this framework is found in its evaluation of the change in a participating individual’s average economic outcome against those individuals who choose not to participate (Cameron and Trivedi, 2005). Individuals included in the non-
participant group are used to form a benchmark to estimate the changes brought about by the treatment.

**Propensity Scores**

In economic studies that use non-experimental data, making treatment participation non-random, and instead reliant upon a vector of observed variables \((X_i)\), propensity score techniques may be able to be utilized (Cameron and Trivedi, 2005). Individuals in the authorized group are directly compared, or matched, to individuals with similar characteristics in the unauthorized group, which allows for the estimation of treatment effects as if a controlled experiment had been conducted. Unfortunately, exact matches do not exist for all authorized individuals.

To correct for these problems, Rosenbaum and Rubin (1983), proposed a propensity score that increases the balance (the similarity between the distribution of \(X_i\) in the control and treatment groups) by estimating a propensity score for each observation and then matching based on that propensity score. The propensity score measures the likelihood that an individual selects into the treatment group, and are typically found using a logistic regression as expressed in equation (10):

\[
P(D_i = 1|X_i) = \frac{e^{F(X_i)}}{1 + e^{F(X_i)}}
\]

where the propensity score \(P(D_i = 1|X_i)\) is the probability that \(D_i\) (the decision to select into authorization) equals 1 given \(X_i\). \(F(X_i)\) is a function of the explanatory variables used in the logistic regression. The propensity score is used to create matches between authorized and
unauthorized individuals to estimate the treatment effects. This study utilizes a logit regression to generate these propensity scores.

In treatment effects estimation, the Conditional Independence Assumption (CIA) is an important assumption. It states that conditional upon observed variables, $X_i$, the potential outcomes are independent of treatment (Cameron and Trivedi, 2005):

$$(11) \quad (Y_{0i}, Y_{1i} \perp D_i | X_i)$$

If it holds, then the matching estimator should perform as expected (Rosenbaum & Rubin, 1983; Heckman, Ichimura, and Todd, 1997; Dehejia and Wahba, 2002; and Smith and Todd, 2005). Presuming this assumption holds, it is understood that treatment status is random conditional upon the vector of observable variables, $X_i$. Thus, matching is a quasi-experimental technique used to replicate actual experimental conditions. It is satisfied if $X_i$ includes all of the variables which affect the selection into treatment and the outcome (Kandilov and Kandilov, 2010).

Using data from the National Agricultural Workers Survey (NAWS), Kandilov and Kandilov (2010) employed propensity score matching techniques$^3$ to compare single male legal permanent residents employed in crop farm work to an appropriate control group of unauthorized workers, in order to analyze the effects of authorization on the wages and benefits. Other studies that have employed the propensity score matching techniques include Dehejia and Wahba (2002) to identify the wage effects of job training programs, and Yasar and Morrison Paul (2008) to analyze the impact of foreign technology transfer on plant productivity. Additionally, Liu and

---

$^3$ As developed by Rubin and Rosenbaum (1983)
Lynch (2011) use propensity score matching methods to examine the effects of agricultural preservation programs and changes in farmland loss.

Using PSM, we estimate three population means:

1. **Average Treatment Effect (ATE):** The expected gain that a randomly selected individual accrues from participating in a program:

   \[
   \alpha_{ATE} = E(\Delta) = E(Y_1) - E(Y_0) \\
   = E(Y_1 | X_i, D = 1) - E(Y_0 | X_i, D = 0)
   \]

   In (12), \(E(\Delta)\) denotes the expected difference between the two different outcomes (\(Y_1\) and \(Y_0\)); \(E(y_1 | X_i, D = 1)\) denotes the outcome (wage rate) for an individual conditional upon a vector of observed variables, \(X_i\), and \(D = 1\) (denoting that the observation has been treated (authorized)), and \(E(y_0 | X_i, D = 0)\) denotes the outcome for the individual conditional upon the same vector of observed variables, \(X_i\), and \(D = 0\) (denoting that the observation is untreated (unauthorized)).

2. **Average Treatment Effect on the Untreated (ATEU):** The effect for non-participants if they would have participated. The ATEU may be useful in future policy decisions regarding the extension of the treatment to excluded groups (Caliendo, 2006). In the context of this study, the ATEU would yield information related to the wage effect that an unauthorized worker could experience if the worker gained legal status. This would also yield information to the grower related to costs of labor if a large-scale legalization program was targeted at the agricultural sector. The ATEU is mathematically expressed as:

   \[
   \alpha_{ATEU} = E(\Delta | D = 0) = E(Y_1 | D = 0) - E(Y_0 | D = 0)
   \]
Where \( E(\Delta|D = 0) \) denotes the expected difference in the outcome (wage rate) given the individual is untreated (unauthorized); \( E(Y_1|D = 0) \); \( E(Y_i|D = 0) \) denotes the expected outcome for the untreated individual if they had decided to participate in the treatment (authorization), and \( E(Y_0|D = 0) \) denotes the outcome for the untreated individual having selected to remain in the untreated status.

3. **Average Treatment Effect on the Treated (ATET):** The average gain from the treatment on those who opt into the treatment (Heckman, Tobias, & Vytlacil, 2001). The ATET is mathematically expressed as:

\[
\alpha_{ATET} = E(\Delta|D = 1) = E(Y_1|D = 1) - E(Y_0|D = 1)
\]

Where \( E(\Delta|D = 1) \) denotes the expected change in the outcome (wage rate) given that the individual is treated (authorized); \( E(Y_1|D = 1) \) denotes the expected outcome for treated individual, and \( E(Y_0|D = 1) \) denotes the expected outcome for the treated individual had they selected to refrain from treatment.

**Selection Bias**

An issue that arises in the calculation for the ATET is that of selection bias. Selection bias poses a challenge and appears in the ATET, in that the ATET includes a hypothetical outcome of lack of treatment for the individuals who actually received the treatment (Caliendo, 2006). If using non-experimental data, as is the case with this study, this outcome will not be the same as the outcome for those individuals who did not receive treatment (i.e., gain legal status). If selection bias occurs as a result of observed factors, linear regression techniques and propensity score matching are appropriate estimation methods. However, when an individual
selects on unobservables, randomization of treatment is difficult to achieve. Selecting on unobservables suggest that unobservable variables drive the individual’s decision to participate in the treatment, and thus nonrandom assignment into treatment occurs and becomes a concern. Instrumental variables are one remedy when faced with issues stemming from selection on unobservables; however, in the context of this study, a valid instrument that is correlated with the decision and uncorrelated with outcome is unavailable. Therefore, we employ a minimum-biased estimation method to account for potential bias due to selection on unobservables.

**Bias-Minimizing Treatment Effects**

To address bias that may occur due to unobserved variables, this study utilizes the minimum-biased estimator (MBE) proposed by Millimet, and Tchernis (2013). These estimators allow for the bias to be minimized when the Conditional Independence Assumption fails – that is, when bias stemming from unobserved factors play a role. Thus the motivation behind using the minimum-biased treatment effects estimation; it addresses issues regarding the efficacy of the CIA and subsequently the potential bias of the propensity score estimates, and deals directly with selection bias arising from unobserved variables.

Assuming individuals select on unobservables, the MBE trims the estimation sample to include only observations with propensity scores that lie within a certain interval. If the CIA holds, the MBE yields unbiased estimations, and works to minimize the bias when the CIA fails (Millimet and Tchernis, 2013). Millimet and Tchernis (2013) propose that by using observations with propensity scores around the Bias-Minimizing Propensity Score (BMPS), \( P^* \), the bias from
the failure of the CIA can be adequately minimized. The ATET, which is the estimator that is subject to selection bias, when estimated using the MBE is expressed as:

\[
\alpha_{\text{ATE,MBE}} = \frac{\sum_{i \in \Omega} \frac{Y_i T_i}{\hat{P}(X_i)} - \sum_{i \in \Omega} \frac{T_i}{\hat{P}(X_i)}}{\sum_{i \in \Omega} \frac{T_i}{\hat{P}(X_i)}} - \frac{\sum_{i \in \Omega} \frac{Y_i (1-T_i)}{1-\hat{P}(X_i)} - \sum_{i \in \Omega} \frac{1-T_i}{1-\hat{P}(X_i)}}{\sum_{i \in \Omega} \frac{1-T_i}{1-\hat{P}(X_i)}}
\]

where \( \Omega = \{i | \hat{P}(X_i) \in C(P^*) \} \) and \( C(P^*) \) symbolizes the neighborhood around \( P^* \) (Millimet & Tchernis 2013). More specifically, \( C(P^*) \) is mathematically expressed as

\[
C(P^*) = \left[ \hat{P}(X_i) \mid \hat{P}(X_i) \in \left( \underline{P}, \overline{P} \right) \right],
\]

where \( \left( \underline{P}, \overline{P} \right) \) is the interval in which a certain percentage of the control and treatment groups’ propensity scores must fall into \( P^* \) (Millimet and Tchernis, 2013). To estimate \( P^* \), \( \Omega \) must be derived. Millimet and Tchernis propose using Heckman’s bivariate normal selection model to estimate \( \Omega \).

The ATET using the MBE is expressed as:

\[
\alpha_{\text{ATE,MBE}}(P^*) = \sum_{i \in \Omega} Y_i T_i - \left[ \frac{\sum_{i \in \Omega} \frac{Y_i (1-T_i) \hat{P}(X_i)}{1-P(X_i)}}{\sum_{i \in \Omega} \frac{(1-T_i) \hat{P}(X_i)}{1-P(X_i)}} \right]
\]

The ATET can be calculated without estimating \( P^* \) (Millimet & Tchernis 2013).

Following Peel (2018), the MBE is used to check the robustness of the traditional propensity-score matching treatment effects estimates, thereby assessing the threat of unobserved

---

4 See Millimet and Tchernis (2013) for detailed mathematical derivations.
selection bias. The CIA is strengthened if the MBE treatment effect estimate is similar and significant, while the presence of an unobserved correlated variable is detected if the MBE treatment effect estimate is still significant, but not similar to PSM estimates (Millimet and Tchernis, 2013; Peel, 2018).

**Limitations and Other Considerations**

A limitation posed by propensity score matching (PSM) is that it typically must utilize a smaller sample size, which therefore reduces the power of the statistical tests conducted. However, unless the PSM technique is employed with an already small sample size, this should not cause any issues in terms of the validity of the PSM treatment effects estimates. In this analysis, authorized (treated) individuals are matched with unauthorized (untreated) individuals using the nearest-neighbor matching technique with replacement. The MBE further trims the sample to include only observations with a propensity score within a certain interval around the BMPS (Millimet and Tchernis, 2013; McCarthy, Millimet, and Tchernis, 2014; Peel, 2018).

Another limitation of PSM, and subsequently the MBE, occurs due to the fact that observations that are not on the common support are discarded. The observed difference in MBE and PSM treatment effects estimates may in part be due to heterogeneity that occurs from individuals’ different responses to selection into treatment (Peel, 2018). However, according to Millimet and Tchernis (2013 p. 988), unobserved bias is still minimized using the BME because it identifies the parameter that can be estimated with least amount of bias.
Results

Characterization of the Sample

The total estimation sample (N=11,958) was drawn from the National Agricultural Workers Survey included on foreign-born, FVH farmworkers from 1989-2014. Of the total sample, 7,851 (65%) observations lacked proper U.S. work authorization, while 4,107 (35%) observations had some type of legal status. Restricting the sample to only include California farmworkers (N=5,927), 2,207 (37%) observations had some type of legal status, while 3,720 (62%) observations did not. Grouping the other NAWS crop regions together, excluding California, yielded a total sample of 6,031 observations; of which, 1,900 (32%) observations had legal status and 4,131 (68%) observations did not.

Nationally, the average authorized foreign-born farmworker was roughly 12 years older than the average unauthorized foreign-born farmworker. In California, the average unauthorized worker was slightly younger than the national average, and the average authorized worker was slightly older than the national average. Foreign-born farmworkers in the ‘all other regions’ grouping, the average unauthorized worker was 32 years old, and the average authorized worker was 44 years old. A higher proportion of authorized workers (nationally, in California, and in all other regions) were married and a higher proportion had English-speaking ability.

Regarding employment, authorized workers earned more their unauthorized counterparts nationally, in California, and in all other regions, with the largest wage differential between the two groups observed in California ($1.15). This larger differential could be due to the higher availability of supervisory positions available in California, 86 percent of which are held by authorized farmworkers. Authorized workers had been in the U.S. longer and had a greater
amount of U.S. farmworker experience. A greater number of unauthorized workers had been hired by a farm labor contractor, as opposed to being hired directly by the grower.

**Logit Specification and Estimation Results**

Table 1 includes descriptions of the variables used in the logistic regression to determine the propensity scores. Variables included in the logit model to predict legal status include: age, experience, experience squared, English-speaking ability, whether the farmworker was hired by a farm labor contractor, whether the farmworker was employed seasonally, the annual number of farmwork weeks, the annual number of weeks spent abroad, the number of years since the farmworker immigrated to the U.S. for the first time, educational attainment, if the farmworker entered the U.S. after 1986, and if the farmworker entered the U.S. after 2001. The selected variables have been used previously in the literature to determine legal status, and there are notable differences between the authorized and unauthorized in terms of means. The last two variables, after 1986 and after 2001, distinguish between two time periods. Following Walters, Emerson and Iwai (2008) the first indicates whether the farm worker first entered the United States to live or work, and is included to reflect authorization through the Special Agricultural Worker (SAW) program that those farmworkers who were already living in the United States and working prior to the 1986 Immigration Reform and Control Act (IRCA) being passed, and the relative difficulty of acquiring legal status since that time. The after 2001 variable reflects farmworker entry to the United States pre and post 2001; in the post-9/11 period following the terror attacks, immigration enforcement (border and interior) was significantly increased.

The model correctly classified 83.93% of observations nationally, 82.83% of observations in California, and 84.84% of observations in all other regions. Figures 1-3 graphically display the distribution of propensity scores pre-matching and post-matching.
Evidenced in figure 3, the distributions of propensity score for the matched sample of unauthorized workers is very similar to the distribution of propensity scores for authorized workers (see figure 2), implying that the two groups have similar characteristics.

The results of the logit specification for the U.S., California, and all other regions are reported in Table 2. Most of the signs on the coefficients are as expected. An immigrant worker is more likely to be authorized to work in the U.S. if the worker is older in age, is proficient with the English language, and is employed seasonally. Additionally, the worker is more likely to be authorized if the worker has resided in the U.S. for a longer period (as captured by the ‘years since immigration’ variable), has more overall U.S. farmwork experience, and has more formal education. On the other hand, a worker is more likely to be unauthorized if the worker has a higher number of annual farmwork weeks and is hired by a FLC (as opposed to being employed directly by the grower). The finding on annual farmwork weeks may reflect that many unauthorized immigrants work a greater number of hours in farm work weekly than authorized immigrant workers, and perhaps because they may not have the same degree of job mobility (in and out of farm work) as their authorized cohorts likely would.

Additionally, the logit results indicate if the worker entered the U.S. for the first time to live or work after 1986 or 2001, they were more likely to be unauthorized. The results are broadly consistent with Walters, Emerson and Iwai (2008). On this particular finding, they concluded this reflected that indicated that immigrant farmworkers following these periods were more likely to be undocumented than otherwise, given the restrictions that were put in place after both of these periods, and heightened enforcement after 2001.

Table 3 presents the estimated propensity-score matching treatment effects parameters at the national level, for California only, and all other regions excluding California. The wage
differences and percentage differences in each region (i.e. national, California, or all other regions) are differences relative to each area’s respective average wage. At the national level, the estimated average treatment on the treated, or the estimated benefit that workers who selected into legal status would experience, on average, is $0.89 per hour. This means that the average benefit from selecting into legal status is associated with a $0.89 increase in the average wage rate. Translating the national wage difference into percentage terms, the average treatment on the treated yielded a 7 percent increase in average wage. The national ATET is broadly comparable to findings from Walters, Emerson, and Iwai (2008) and Kandilov and Kandilov (2010).

Additionally, overall, our findings are consistent with those of previous work that have assessed the wage implications of legal status (Taylor, 1992; Isé and Perloff, 1995; Kossoudji and Cobb-Clark, 2002; Iwai, Emerson, and Walters, 2006; Walters, Emerson, and Iwai, 2008; Kandilov and Kandilov, 2010). In California, the average treatment on the treated is higher than the national effect, with workers who selected into legal status experiencing a $1.09 per hour increase in average wages, which translates into a 9 percent increase in wage rates of California foreign-born farmworkers. Matching performed using observations from all other regions yielded an average treatment effect of $0.65 per hour, or roughly a 5 percent increase in average wage rates of farmworkers found in regions other than California. The average treatment effect on the treated is significant at the national level, for workers in California, and for workers grouped in all other regions excluding California.

The average treatment effect on the untreated (the effect that untreated workers would experience if they selected into legal status), is only significant at the national level. The average treatment effect on the untreated at a national level is an estimated $0.29 per hour, or a 2 percent, increase in average wage rates. For untreated workers in California, there is an estimated $0.49
per hour, or a 5 percent, increase in average California wage rates. In the group encompassing workers in all other regions, there is an estimated $0.02 per hour, or a 0.7%, increase in average wage rates.

The estimated average treatment effect (the wage effect that a worker who randomly selected into legal status would experience) is positive and significant nationally, as well as in California and for the group of all other regions. The national average treatment effect is estimated to be a $0.49, or a 4 percent, increase in wages. For California workers, the average treatment effect is estimated to be higher than the national effect, with the randomly selected worker experiencing a $0.71, or 6 percent, increase in wage rates due to receiving the treatment. In the all other regions grouping, the average worker could expect an estimated $0.22, or 2 percent, increase in wage rates.

Results from the minimum-biased estimation technique are reported and compared to the treatment effects estimates calculated from propensity score matching techniques in Table 4. The results are presented in terms of percentage differences between the wages of the treated (authorized) and untreated (unauthorized), and are interpreted as suggested by Peel (2018).

Comparison of the PSM and MBE methods indicate, in general, that immigrant workers who select into legal status for US farm work earn higher wages than the average immigrant worker who gains legal status, and more than the average undocumented immigrant worker. At the national level, the PSM reports that those immigrant farmworkers who selected into legal status ($ATET_{PSM}$) earn wages that are 3% higher than the average immigrant worker who randomly gained legal status ($ATE_{PSM}$), and 5% higher than those who remained undocumented ($ATU_{PSM}$). This gap is shown to narrow to 2% in both of the latter cases, when the MBE is used.
to account for selection bias that arises from unobserved sources – which may be peculiar
characteristics about the workers or labor market conditions that are unknown to the researcher.

A similar direction of effect is apparent for California, and the magnitude of the gap is the same for the $\text{ATET}_{\text{PSM}}$ and the $\text{ATE}_{\text{PSM}}$. However, the gap between the estimates for the $\text{ATET}_{\text{MBE}}$ and the $\text{ATU}_{\text{MBE}}$ is much lower (1%) than that the $\text{ATET}_{\text{PSM}}$ and $\text{ATU}_{\text{PSM}}$ (4%). This suggests that the PSM would largely overstate the effect of legal status on the wages for authorized versus unauthorized workers. Also interesting, while the $\text{ATET}_{\text{PSM}}$ and $\text{ATE}_{\text{PSM}}$ for the combined regions show the same magnitude of difference, the largest gap of the three regions modeled in the analysis is apparent: the $\text{ATET}_{\text{PSM}}$ is 4.3% higher than the $\text{ATU}_{\text{PSM}}$. However, this effect disappears when for the MB estimates of these population means. This appears to underscore the importance of recognizing that there are likely unobserved characteristics driving the selection bias, which would result in the wage implications of legal status being overestimated, if ignored.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age of farm worker at time of interview</td>
</tr>
<tr>
<td>U.S. Farmwork Experience</td>
<td>Years of U.S. Farmwork</td>
</tr>
<tr>
<td>U.S. Farmwork Experience, sq.</td>
<td>Years of U.S Farmwork squared</td>
</tr>
<tr>
<td>English-Speaking Ability</td>
<td>=1 if worker has the ability to speak English =0 if worker does not</td>
</tr>
<tr>
<td>Employed by Farm Labor Contractor</td>
<td>=1 if worker is employed by FLC =0 if worker is not</td>
</tr>
<tr>
<td>Seasonal Labor</td>
<td>=1 if worker is classified as seasonal labor =0 if worker is not</td>
</tr>
<tr>
<td>Annual Farmwork Weeks</td>
<td>Number of weeks worked on-farm during the last year</td>
</tr>
<tr>
<td>Annual Weeks Spent Abroad</td>
<td>Number of weeks spent outside of the U.S. during the last year</td>
</tr>
<tr>
<td>Years Since Immigration</td>
<td>Number of years since worker entered the U.S. to live or work for the first time</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>Number of years of traditional schooling</td>
</tr>
<tr>
<td>Entered U.S. after 1986</td>
<td>=1 if worker entered the U.S. for the first time after 1986 =0 if entered before 1986</td>
</tr>
<tr>
<td>Entered U.S. after 2001</td>
<td>=1 if worker entered the U.S. for the first time after 2001 =0 if entered before 2001</td>
</tr>
<tr>
<td>Variables</td>
<td>National</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Age</td>
<td>0.0207***</td>
</tr>
<tr>
<td></td>
<td>(0.00341)</td>
</tr>
<tr>
<td>U.S. Farmwork Experience</td>
<td>0.0118</td>
</tr>
<tr>
<td></td>
<td>(0.0127)</td>
</tr>
<tr>
<td>U.S. Farmwork Experience, sq.</td>
<td>0.000853***</td>
</tr>
<tr>
<td></td>
<td>(0.000357)</td>
</tr>
<tr>
<td>English-Speaking Ability</td>
<td>0.468***</td>
</tr>
<tr>
<td></td>
<td>(0.0584)</td>
</tr>
<tr>
<td>Employed by FLC</td>
<td>-0.324***</td>
</tr>
<tr>
<td></td>
<td>(0.0792)</td>
</tr>
<tr>
<td>Seasonal Labor</td>
<td>0.156***</td>
</tr>
<tr>
<td></td>
<td>(0.0577)</td>
</tr>
<tr>
<td>Annual Farmwork Weeks</td>
<td>-0.0168***</td>
</tr>
<tr>
<td></td>
<td>(0.00242)</td>
</tr>
<tr>
<td>Annual Weeks Spent Abroad</td>
<td>-0.00662</td>
</tr>
<tr>
<td></td>
<td>(0.00416)</td>
</tr>
<tr>
<td>Years Since Immigration</td>
<td>0.0561***</td>
</tr>
<tr>
<td></td>
<td>(0.00725)</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>0.0755***</td>
</tr>
<tr>
<td></td>
<td>(0.00876)</td>
</tr>
<tr>
<td>Entered U.S. after 1986</td>
<td>-1.525***</td>
</tr>
<tr>
<td></td>
<td>(0.0940)</td>
</tr>
<tr>
<td>Entered U.S. after 2001</td>
<td>-0.327***</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.727***</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
</tr>
</tbody>
</table>

Observations: 11,956, 5,927, 6,031

*Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*
Figure 1. Distribution of Propensity Scores – All Undocumented Workers

![Figure 1. Distribution of Propensity Scores – All Undocumented Workers](image1)

Figure 2. Distribution of Propensity Scores – All Authorized Workers

![Figure 2. Distribution of Propensity Scores – All Authorized Workers](image2)

Figure 3. Propensity Score Distribution – Matched Undocumented Workers

![Figure 3. Propensity Score Distribution – Matched Undocumented Workers](image3)
Table 3. Propensity Score Matching Treatment Effects Estimation

<table>
<thead>
<tr>
<th></th>
<th>National</th>
<th>California</th>
<th>All Other Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wage %</td>
<td>Wage %</td>
<td>Wage %</td>
</tr>
<tr>
<td>ATET</td>
<td>$0.89***</td>
<td>$1.09***</td>
<td>$0.65***</td>
</tr>
<tr>
<td></td>
<td>7%***</td>
<td>9%***</td>
<td>5%***</td>
</tr>
<tr>
<td>ATU</td>
<td>$0.29*</td>
<td>$0.49</td>
<td>$0.02</td>
</tr>
<tr>
<td></td>
<td>2%*</td>
<td>5%</td>
<td>0.7%</td>
</tr>
<tr>
<td>ATE</td>
<td>$0.49***</td>
<td>$0.71***</td>
<td>$0.22*</td>
</tr>
<tr>
<td></td>
<td>4%***</td>
<td>6%***</td>
<td>2%*</td>
</tr>
<tr>
<td>Sorting Gains</td>
<td>$0.40</td>
<td>$0.39</td>
<td>$0.43</td>
</tr>
<tr>
<td>Matching</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Distance</td>
<td>0.0003</td>
<td>0.0009</td>
<td>0.0007</td>
</tr>
<tr>
<td>N</td>
<td>11,958</td>
<td>5,927</td>
<td>6,031</td>
</tr>
</tbody>
</table>

Asterisks denote significance - ***p<.01, **p<.05, *p<.1

1 Sorting gains = ATET – ATE
2 Matching distance = abs (propensity score – propensity score of nearest neighbor)

Table 4. Bias-Minimizing Treatment Effects Estimation

<table>
<thead>
<tr>
<th></th>
<th>National</th>
<th>California</th>
<th>All Other Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSM</td>
<td>MBE</td>
<td>PSM</td>
</tr>
<tr>
<td>ATET</td>
<td>7%***</td>
<td>5%***</td>
<td>9%***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATU</td>
<td>2%*</td>
<td>3%**</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>4%***</td>
<td>3%***</td>
<td>6%***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>11,958</td>
<td>5,927</td>
<td>6,031</td>
</tr>
</tbody>
</table>

Asterisks denote significance - ***p<.01, **p<.05, *p<.1
Discussion and Concluding Remarks

The purpose of this study was to analyze the wage implications of legalization on foreign-born U.S. crop farm worker wages at the national level, in California, and in a grouping of the other five NAWS crop regions. California Three population means were estimated: (1) the average treatment effect, (2) the average treatment effect on the treated, and (3) the average treatment effect on the untreated. This study contributes to the literature in two ways: first, it assesses the potential wage outcomes on the national relative to regional levels. This was not assessed previously by past studies. Second, this study uses the minimum-biased estimator (MBE) proposed by Millimet and Tchernis (2013) as a robustness check for unobserved selection-bias. To the authors’ knowledge, the minimum-biased estimator has not been applied in the context of this particular research problem. Previous work has acknowledged selection bias that arises from observed sources, but has not explicitly addressed when it arises from unobserved sources.

Due to the non-randomization of the treatment (legal status), an individual either selects into treatment based on observable characteristics or selects into treatment based on unobservable characteristics or based on a mixture of both. This study employs treatment effects estimation methods under the assumption that farmworkers select into legalization based both on observables (via propensity score matching) and unobservables (via minimum-biased estimation). The strength of the estimates provided through propensity score matching rests on the conditional independence assumption, which signifies that selection into treatment is random conditional upon a specified vector of covariates. If the conditional independence assumption fails to hold, propensity score estimates become subject to selection bias. To detect the presence of selection on unobservables, and to provide the treatment effect estimate with the least bias
(assuming the conditional independence assumption fails to hold), minimum-biased estimators are employed. Propensity scores were calculated via a logistical regression and authorized observations with were matched with unauthorized counterparts based on similar covariates.

Results from the matching show that there are statistically significant positive wage effects at the national level as well as for California and the rest of the United States. California is singled out for comparison as a major specialty crop producing state\(^5\) with high labor intensity relative to the rest of the United States. Further, California has significant activity by farm labor contractors (FLCs) in its farm labor market, relative to other states. California, as a single crop state/region in the NAWS data, has a large immigrant workforce, much of which is unauthorized for US employment. The average treatment effect on the treated is largest in California at $1.09 more for authorized workers relative to unauthorized, while the wage effect for all other regions is $0.65. The national wage effect of authorization is $0.89. The average treatment effect (the wage effect that a randomly selected worker would experience if the work moved from unauthorized to authorized) is positive and significant nationally ($0.49) as well as in California ($0.71) and all other regions ($0.22). The average treatment effect on the untreated is significant at the 5% level nationally ($0.29) and at the 10% level for California ($0.49), but insignificant for all other regions. As discussed in Caliendo (2006), the average treatment effect on the untreated highlights the potential policy implications for the future. These results are a signal of what could happen if policies were put in place to make it easier for unauthorized workers to transition to authorized status or if another large-scale amnesty program was targeted toward the agricultural industry.

\(^5\) California is designated as a distinct crop region in the NAWS data.
California tends to stand out relative to the rest of the country in terms of the impact of authorization on farm workers. The treatment effect is higher in California than the rest of the country for both the ATET and the ATU. This is important because 44% of all foreign farm workers are in California and California supplies roughly 50% of the nation’s fresh fruits and vegetables (Martin, 2014). The political climate is also different in California relative to much of the rest of the country. Immigration laws have been enforced more leniently in California and state-level laws tend to be more favorable for undocumented workers in California relative to much of the rest of the country. This is widely documented in various media reports (Ramakrishnan and Colbern, 2015; Garcias, 2017; FindLaw, 2018). Clearly, this may create a unique environment for farm workers in the state, in that it could potentially be easier to gain legal status.

The order of the magnitude of the estimates (ATET > ATE > ATU) indicates positive sorting regarding the gains from obtaining legal status. Foreign-born farm workers who had taken deliberate actions to obtain legal status gained more from it than the average, randomly-selected worker, and more than those that were unauthorized. This suggests that farm workers may be weighing the benefits of obtaining work authorization with the costs of becoming authorized. Arguably, those who stand to gain the most may tend to take the steps necessary to become authorized while those who are likely to gain the least may opt to remain unauthorized. In the latter case, it is possible for there to be unique factors that may disqualify the immigrant worker, that are unobserved by the researcher. The process to become legalized is known to be time and effort intensive, as well as expensive and requires a knowledge of how to navigate the process of obtaining some type of legal status (i.e. green card, legal permanent resident, permanent U.S. citizen). For the average unskilled and undocumented farm worker, legal status
may not be an option. Further, the average undocumented worker may also perceive it as risky, in that it could potentially increase detection and deportation risks. Clearly, these are factors that may not be easily documented on any survey or observed by researchers, although they could impact undocumented workers.

The minimum-biased estimator provided a higher (compared to the ATET provided by propensity score matching), and still significant, estimate of the ATET. This finding implies that propensity score matching estimates are underestimated (due to the presence of an unobserved correlated variable) and therefore the wage effects provided by the propensity score matching method are understated. The ATE and the ATU estimates provided by minimum-biased estimation were also higher, and still significant, suggesting that the ATE and ATU provided by propensity score matching methods are overstated due to the presence of an unobserved correlated variable. After comparing the estimates provided by the different estimation methods, it becomes apparent that immigrant farm workers individuals likely select into legal status based on unobserved factors, which lowers confidence in the strength of the conditional independence assumption. Both sources of selection bias would need to be acknowledged.

Focusing on the minimum-biased (MBE) estimates relative to the propensity score (PSM) estimates, PSM estimates were higher than MB estimates. This suggests the PSM method overstates the average treatment effect on the treated (ATET) (nationally, in California, and in all other regions), the average treatment effect on the untreated (California), and the average treatment effect (nationally and California), due to unobserved variables that are correlated with both the outcome and treatment. If the variable(s) could be included in the analysis, the PSM estimate would likely be lower.
The PSM estimate of the average treatment effect on to the untreated (ATU) (national), and the average treatment effect (all other regions) was lower than the MBE suggesting that the PSM estimate is understated due to unobserved correlated variable(s). In the case of a PSM overstatement of the effect, the unobserved variable is positively correlated with both the outcome variable and the treatment variable. Conversely, in the case of a PSM understatement of the effect, the unobserved variable is positively correlated with the outcome variable and negatively correlated with the treatment variable (Peel, 2018). The awareness of the presence of selection bias due to unobservable factors shows the importance of utilizing methods that correct for such bias.

Beginning with the average treatment effect on the treated, the MBE estimate is lower than the PSM in all three regional specifications. Like the PSM ATET estimate, the MBE ATET estimate is significant, signifying that the ATET estimate provided by PSM is overestimated due to the presence of an unobserved correlated variable. The minimum-biased estimate for the average treatment effect on the untreated is higher than the PSM ATU estimates, and significant in California, unlike the PSM ATU estimate for California. This implies that the PSM ATU estimates are underestimated nationally and in the group containing all other regions, whereas PSM ATU estimates are overestimated in California.

Regarding the average treatment effect, the MBE ATE estimates are all significant, however, the MBE ATE estimates are lower nationally and in California, implying that the PSM ATE estimates are potentially overestimated due to the presence of an unobserved correlated variable. In the all other regions grouping, the MBE ATE estimate is higher, suggesting that the PSM ATE estimate could be underestimated due to an unobserved correlated variable.
In sum, these results appear to suggest that studies that do not address this particular source of bias would overestimate that wage implications of legal status, and, by extension, the potential hired labor cost increases.

Evaluating the precise economic implications of legalization for foreign born workers in the US economy can be challenging considering the myriad data collection and estimation issues that arise, not to mention the dynamic political environment in which legislation on the state and national levels are made. Foreign-born individuals who seek employment would make such decisions based on a variety of factors, only some of which some are actually observable. There may be certain additional characteristics, behaviors and attitudes of immigrant workers, and/or conditions of the work environment, the labor market and the immigration policy environment that may need to be taken into account. The regional analysis conducted in this study was restricted to two regions – California and all other regions – due to sample size restrictions during the minimum-biased estimation. Future research could focus on addressing this issue and, if successful, analyze the effects across more regions, to ascertain how these may differ accordingly, particularly in light of the implementation of various immigration initiatives at the state level. Additionally, future research could attempt to assess the regional effects using other comprehensive data sets, if available. Although such analyses may be sensitive to the methods of estimation, they would contribute to greater understanding of potential impacts of immigration reform, especially if it were to affect the supply of workers and hired labor costs. This is crucial for labor intensive sectors in US agriculture, given the ramifications that it could have for agricultural investment and production decisions and producer livelihoods.
References


