The Mario Einaudi Center
for International Studies

Working Paper Series

How Important is Selection?

Experimental Versus Non-experimental Measures of the Income Gains from Migration

by

David McKenzie, John Gibson and Steven Stillman

Paper No. 04-06

April 2006

**The Mario Einaudi Center
for International Studies**

170 Uris Hall
Cornell University
t. 607-255-6370
f. 607-254-5000
Einaudi_Center@is.cornell.edu
www.einaudi.cornell.edu

How Important is Selection? Experimental Versus Non-experimental Measures of the Income Gains from Migration[1]

by

David McKenzie, John Gibson and Steven Stillman

Abstract

Measuring the gain in income from migration is complicated by non-random selection of migrants from the general population, making it hard to obtain an appropriate comparison group of non-migrants. This paper uses a migrant lottery to overcome this problem, providing an experimental measure of the income gains from migration. New Zealand allows a quota of Tongans to immigrate each year with a lottery used to choose amongst the excess number of applicants. A unique survey conducted by the authors in these two countries allows experimental estimates of the income gains from migration to be obtained by comparing the incomes of migrants to those who applied to migrate, but whose names were not drawn in the lottery, after allowing for the effect of noncompliance among some of those whose names were drawn. We also conducted a survey of individuals who did not apply for the lottery. Comparing this non-applicant group to the migrants enables assessment of the degree to which non-experimental methods can provide an unbiased estimate of the income gains from migration. We find evidence of migrants being positively selected in terms of both observed and unobserved skills. As a result, non-experimental methods are found to overstate the gains from migration, by 9 to 82 percent. A good instrumental variable works best, while difference-in-differences and bias-adjusted propensity-score matching also perform comparatively well.

Keywords: Migration, Selection, Natural Experiment
JEL codes: J61, F22, C21

Contact Information

David McKenzie, MSN MC3-300, The World Bank, 1818 H Street N.W., Washington D.C. 20433, USA. E-mail: dmckenzie@worldbank.org,
Tel: (202) 458-9332, Fax (202) 522-3518.

[1] We thank the Government of the Kingdom of Tonga for permission to conduct the survey there, the New Zealand Immigration Service for providing the sampling frame, Halahingano Rohorua and her assistants for excellent work conducting the survey, and most especially the survey respondents. Alan de Brauw, Chirok Han, Martin Ravallion, Ed Vytlacil and participants at seminars at Columbia University, NEUDC, NZESG, NZIS, the University of Canterbury, and the World Bank provided helpful comments. Financial support from the World Bank, Stanford University, the Waikato Management School and Marsden Fund grant UOC0504 is gratefully acknowledged. The views expressed here are those of the authors alone and do not necessarily reflect the opinions of the World Bank, the New Zealand Immigration Service, or the Government of Tonga.

1. Introduction.

Is migration a good investment? To determine the income gains from migration, one must compare the earnings of the migrant to what they would have earned in their home country. The latter is unobserved, and is usually proxied by the earnings of stayers of a similar age and education to the migrant. This approach is not very convincing because if the two groups are really the same, they should have the same migratory behaviour (Lalonde and Topel, 1997). Simple comparisons of movers and stayers are therefore likely to be misleading, as differences in outcomes may just reflect unobserved differences in ability, skills, and motivation, rather than the act of moving itself. Recognizing this difficulty, economists often use statistical corrections for non-random selection when modelling outcomes for migrants (Robinson and Tomes, 1982). However, there is some doubt about the assumptions behind these statistical remedies for selectivity in non-experimental data (Deaton, 1997), especially when the odds of migrating are very low (Hartog and Winkelmann, 2003). These doubts persist because it is hard to know how well these remedies compare with the ideal of a randomized experiment.

The research reported here uses a unique random selection mechanism to overcome the interpretation difficulties posed by the non-random selection of migrants, and then compares experimental estimates of the gains from migration to results obtained using non-experimental estimation methods. The random selection mechanism we use is based on the Pacific Access Category (PAC) under New Zealand's immigration policy. The PAC allows an annual quota of Tongans to migrate to New Zealand without going through the usual migration categories used for groups such as skilled migrants and business investors. Many more applications are received than the quota allows, so a lottery is used by the New Zealand Immigration Service to randomly select from amongst the registrations. A survey administered by the authors was used to collect data on winners and losers in this lottery. Thus, we have a group of migrants and a comparison group who are similar to the migrants, but remain in Tonga only because they were not successful in the lottery.

By comparing the lottery winners and losers, we are able to obtain the only known experimental measure of the gain in income from migration. As not all individuals whose names were selected in the lottery had migrated by the time of our survey, this estimate accounts for non-compliance to the "treatment" of migration. We therefore consider both the intention-to-treat effect, which is the impact on expected income of having a winning ballot in the PAC lottery, and the average treatment effect on the treated, which is the average impact of migrating for individuals who migrate after winning the lottery. We estimate that there is an 88% increase in expected income from winning the lottery, and a 263% increase in income from migrating.

In addition to winners and losers in the PAC lottery, we also surveyed individuals who did not apply for the lottery. We use this sample of non-applicants along with the migrant sample to obtain non-experimental estimates of the income gains from migration. Five popular non-experimental methods for dealing with selectivity are considered: a single difference estimator which compares post-migration income to pre-migration income; OLS regression estimates which assume selection on observables; difference-in-differences regression estimation; propensity-score matching; and instrumental variables using the pre-existing migrant network and the pre-migration distance from the office in Tonga where ballot registrations are deposited as instruments. Each of these methods is found to overstate the gain in income from migration compared to the experimental estimate. Instrumental variables using a good instrument (pre-migration distance) performs best, only overstating the gains by 9%. The single-difference estimator overstates the gains by 25%, while difference-in-differences overstates the gains by 20%. Propensity-score matching overstates the gains by 19-33%, doing better when past income is included as a control and when the bias-adjusted methods of Abadie and Imbens (2005) are used. OLS overstates the gains by 31%, while a poor instrument (the size of the migrant network) overstates the gains by 82%, which is almost as large as the bias in the simple cross-country comparison of GDP per capita (100% overstatement).

The existing empirical literature on migrant selectivity focuses exclusively on observable measures of skills, such as education. For example, Chiquiar and Hanson (2005) find Mexican immigrants to the United States to be positively selected in terms of education and other observable skills. This contrasts with the model of Borjas (1987), which predicts that individuals moving from a country with a less equal income distribution to one with a more equal income distribution will tend to be negatively selected from their home country distribution. The Gini of weekly earnings from wage, salary and self-employment work in Tonga is 0.338, compared to a Gini of 0.374 in New Zealand[2], so Borjas's model would predict positive selection from Tonga. The overstatement of the

[2] Tonga Gini calculated from our sample of workers in non-migrant households; Gini for New Zealand calculated from the 2002 New Zealand Income Survey.

income gains from migration obtained from the non-experimental methods is consistent with this theory, if migrants from Tonga are positively selected in terms of unobserved ability and skills, conditional on their observed characteristics. We examine selection directly by looking at pre-migration earnings, and do find migrants to be positively selected in terms of unobserved characteristics, with most of this occurring through selection into the lottery, rather than through selective compliance conditional on winning the lottery.

This paper also contributes to the literature started by the influential work of Lalonde (1986), which attempts to assess the ability of non-experimental estimators to obtain estimates similar to experimental results. To date, this literature has concentrated on a small number of labor market training programs. After Lalonde's initial pessimistic assessment of non-experimental measures, there has been much recent debate as to the ability of propensity-score matching methods to obtain better results (e.g. Heckman, Ichimura and Todd, 1997; Dehejia and Wahba 2002; Smith and Todd 2005; Dehejia 2005). The migration example we consider here offers many of the features identified by these studies as conducive to more accurate non-experimental estimation. The non-migrant control group were administered the same survey instrument as the migrants, including retrospective earnings information, and live in the same villages and work in the same labor markets. Unlike in many labor program settings, there is no substitution bias, as the ability of the controls to migrate other than through the program we consider is severely limited. Moreover, the size of the "treatment" considered here is large and strongly significant. This contrasts with a treatment effect of a 29% increase in earnings in Lalonde's NSW male sample, which only had a t-statistic of 1.82. Even with these favorable conditions, the non-experimental estimators still overstate the income gains. However, we find that the more recent refinements of propensity-score matching do enable more precision, and provide point estimates which are not statistically different from the experimental estimator.

The remainder of this paper is structured as follows. Section 2 describes the immigration process used as the natural experiment and the sampling method and data from the Pacific Island-New Zealand Migration Study. Section 3 constructs the experimental estimates. Section 4 estimates five different types of non-experimental estimates. Section 5 looks directly at selection, Section 6 considers cost-of-living adjustments and Section 7 concludes.

2. The Pacific Access Category and PINZMS Data

The natural experiment we use is based on the Pacific Access Category (PAC) under New Zealand's immigration policy. The PAC was established in 2001 and allows an annual quota of 250 Tongans to migrate to New Zealand without going through the usual migration categories used for groups such as skilled migrants and business investors. Specifically, any Tongan citizens aged between 18 and 45, who meet certain English, health and character requirements, can register to migrate to New Zealand.[3] Many more applications are received than the quota allows, so a ballot is used by the New Zealand Immigration Service (NZIS) to randomly select from amongst the registrations. The

[3] The person who registers is a Principal Applicant. If they are successful, their immediate family (spouse and children under age 18) can also apply to migrate as Secondary Applicants. The quota of 250 applies to the total of Primary and Secondary Applicants, and corresponds to about 70 migrant households.

probability of success in the ballot is approximately 10 percent. Thus, we have a group of migrants and a comparison group who are similar to the migrants, but remain in Tonga only because they were not successful in the lottery. Once their ballot is selected in the lottery, applicants must provide a valid job offer in New Zealand within six months in order to have their application to migrate approved and be allowed to migrate.

The data used here are from the Tongan component of the Pacific Island-New Zealand Migration Survey (PINZMS), a comprehensive household survey designed to measure multiple aspects of the migration process. Questions on household demographics, education, labor supply, income, asset ownership and food consumption, were based where possible on the most widely used surveys in New Zealand and the Pacific Islands to enhance comparability. The survey design and enumeration, which was overseen by the authors in the first half of 2005, covered random samples of four groups: (i) Tongan migrants to New Zealand, who were successful participants in the 2002/03 and 2003/04 PAC lotteries, (ii) successful participants from the same lotteries who were still in Tonga, either because their application for New Zealand residence was not approved (typically because of lack of a suitable job offer) or was still being processed, (iii) unsuccessful participants from the same lotteries who were still in Tonga, and (iv) a group of non-applicants in Tonga.

The initial sample frame for groups (i) and (ii) was a list of the names and addresses of the 278 (out of almost 3000 applicants) successful participants in the 2002/03 and

2003/04 migration lotteries.[4] Approximately 100 of these successful ballots had been approved for residence in New Zealand by the time of the survey, although some of those families had not yet moved to New Zealand. We managed to locate 65 of the families that had migrated, giving a sampling rate of over 70 percent. A variety of tracking methods were used to locate these families including contacting their family back in Tonga and using key informants in churches and other community groups. It was easier to draw a random sample of 55 of the successful ballots that had not yet migrated, because the NZIS records included postal and home addresses and telephone numbers in Tonga. This non-migrant group includes those whose applications were rejected and those whose applications were still being processed. We use the actual number of accepted and rejected applications to weight our sample.

The initial sample frame for the unsuccessful ballots in the 2002/03 and 2003/04 lotteries (group (iii)) was a list of names and addresses provided by the NZIS. The details for this group were less informative than those for the successful ballots. Only a postal address was supplied and there were no telephone numbers. Thus, it was not possible to determine whereabouts in Tonga those with unsuccessful ballots lived. Moreover, many of the postal addresses were either for immigration agents, or were outside of Tonga (especially in New Zealand). We used two strategies to derive a sample of 78 unsuccessful ballots from this information: first, as part of our survey of the migrants in New Zealand we had obtained details about the location of remaining family (almost 60 percent of migrants still had family occupying their previous dwelling in Tonga). We

[4] This was supplied under a contractual arrangement with the New Zealand Immigration Service, with strict procedures used to maintain the confidentiality of participants.

used this information to draw a sample of unsuccessful ballots from the same villages (implicitly using the village of residence when the applicant entered the ballot as a stratifying variable). We also used the Tongan telephone directory to find contact details for people included in the list of names supplied by NZIS. To overcome concerns that this would bias the sample to more accessible areas around the capital city of Nuku'alofa, who are more likely to have telephones, we deliberately included in the sample households from two of the outer Islands, of Vava'u and 'Eua.

Table 1 examines how random the sample we have is by comparing means of ex-ante characteristics for lottery winners and lottery losers among the principal applicants in our sample. The point estimates of the means are similar in magnitude for the two groups and we can not reject equality of means for any of the variables. This is as would be expected with the random selection of ballots among applicants in the Pacific Access Category.

The sample of non-applicants was obtained by selecting 60 households, with at least one member aged 18 to 45, in either the same villages that the migrants had been living in prior to migrating or in the same villages that unsuccessful ballots were found in. An initial screening question was used to check that no-one in the household had previously applied for the migration lottery. Data on employment, income, and demographics was collected on all members of these households. Additional questions on the reasons for not applying, the size of the family networks in New Zealand, and expectations, were asked of the oldest member aged 25-35 in the household, or of the oldest member aged 18-45 if no one was aged 25-35. We will refer to this group of individuals which received the extended questions as the group of pseudo-applicants.

Table 2 presents the proportion employed, mean hours worked, and mean work income among the different groups in our sample. The mean weekly income from work among migrants is NZ$425, compared to $81-104 for applicants for the Pacific Access Category (PAC) lottery who did not migrate, and $41 among all individuals aged 18 to 45 in non-applicant households.[5] A t-test of equality of means strongly rejects the null hypothesis of equality of migrant income with any of the other groups. The point estimates suggest that migrants are more likely to be employed than non-migrants, and work slightly longer hours. However, these differences are not significant given our sample size.

3. Experimental estimates of the income gain from migration

3.1. Estimating treatment effects using experimental data

This paper focuses on estimating the impact of migration to New Zealand on the income of Tongans. To determine the income gains from migration, one must compare the earnings of the migrant to what they would have earned in their home country had they not migrated. Typically, it is not possible to readily identify this unobserved counterfactual outcome. However, the PAC lottery system, by randomly denying eager migrants the right to move to New Zealand, creates a control group of individuals that should have the same outcomes as what the migrants would have had if they had not moved. In our application, a comparison of mean income for lottery winners who migrate and lottery losers can be used to obtain an experimental measure of the gain in

[5] At the time of the survey, NZ$1=US$0.72.

income from migration. This simple comparison of means at the bottom of Table 2 shows a \$320 increase in weekly work income from migrating.

As discussed in Heckman et. al. (2000), this simple experimental estimator of the treatment effect on the treated (SEE-TT) is biased if control group members substitute for the treatment with a similar program or if treatment group members dropout of the experiment. In our application, *substitution* bias will occur if PAC applicants who are not drawn in the lottery migrate to New Zealand through an alternative visa category such as the family or skills category or migrate to another country and *dropout* bias will occur if PAC applicants whose name are drawn in the lottery fail to migrate to New Zealand. We do not believe that substitution bias is of serious concern in our study, as individuals with the ability to migrate via other arrangements will likely have done so previously given the low odds of winning the PAC lottery[6]. However, as shown in Table 2, dropout bias is a more relevant concern; only one-third of lottery winning principal applicants had migrated to New Zealand at the time of our survey. A number of the other individuals are in the process of moving, but the majority of those not migrating are unable to move due to the lack of a valid job offer in New Zealand.

The impact of dropout bias on the SEE-TT of the gain in income from migration estimated above can be illustrated by writing the income of applicant *i* as:

$$\text{Income}_i = \alpha + \beta^*\text{BallotSuccess}_i + \nu_i, \text{ where } E(\nu_i)=0, \qquad (1)$$

[6] We did not come across any incidences where remaining family members told us that the unsuccessful applicant had migrated overseas during our fieldwork.

$BallotSuccess_i$ is a dummy variable taking the value one if the PAC applicant's ballot is drawn in the lottery and zero if it is not drawn, and alternatively as:

$$Income_i = \mu + \lambda * Migrate_i + \varepsilon_i, \text{ where } E(\varepsilon_i) = 0, \tag{2}$$

where $Migrate_i$ is a dummy variable taking the value one if person *i* migrates and zero otherwise, and λ is the average treatment effect on the treated.

The SEE-TT of the gain in income from migration is calculated as the difference in mean income between lottery winners who migrate and unsuccessful ballots:

$$SEE\text{-}TT = E[Income_i \mid Migrate_i=1] - E[Income_i \mid BallotSuccess_i=0] \tag{3}$$

However, from equation (2), we can see that:

$$SEE\text{-}TT = \lambda + E[\varepsilon_i \mid Migrate_i=1] - E[\varepsilon_i \mid BallotSuccess_i=0] \tag{4}$$

Thus, the SEE-TT will only be an unbiased estimate of λ if the last two terms in equation (4) sum to zero. Because ballot success is determined randomly via a lottery we can replace $E(\varepsilon_i|BallotSuccess=0)$ with $E(\varepsilon_i|BallotSuccess=1)$ and rewrite (4) to show that the SEE-TT is an unbiased estimate of the treatment effect on the treated if and only if:

$$E[\varepsilon_i \mid Migrate_i=1] = E[\varepsilon_i \mid BallotSuccess_i=1]. \tag{5}$$

That is, this simple estimator will give a consistent estimate of the income gains from migration if and only if there is no selection as to who migrates among those successful in the lottery. This condition does not seem likely to hold, and in this case estimating the impact of migration requires comparison of other groups.

3.2. Intention-to-treat effect

Experimental data, in the presence of substitution and dropout bias, can identify the mean impact of a program (eg. winning the lottery) on outcomes (eg. income for PAC applicants), also known as the intention-to-treat effect (ITT).[7] This estimator, β in equation (1), is unbiased because randomisation insures that $E(\nu_i \mid BallotSuccess_i=1)$ equals $E(\nu_i \mid BallotSuccess_i=0)$, and can be computed by comparing the mean income for ballot winners to that for ballot losers. As shown at the bottom of Table 2, on average, winning the PAC lottery is estimated to increase weekly income by $91.

While the results in Table 1 show that the lottery did indeed achieve reasonably comparable groups, the small size of our sample may have resulted in some differences between successful and unsuccessful ballots. To improve the efficiency of our ITT estimate, we re-estimate β using an ordinary least squares (OLS) regression model described in equation (6) to add control variables for the observable pre-existing characteristics of the two groups:

$$Income_i = \alpha + \beta * BallotSuccess_i + \delta' X_i + \omega_i \qquad (6)$$

Column 1 of Table 3 first estimates this regression with no controls, repeating the estimate of $91 obtained as the difference in means. In Column 2 we add a set of controls for pre-existing characteristics of applicants. These include standard wage equation variables, such as age, sex, marital status, and years of education. In addition, we include

[7] The terminology *Intent-to-treat* comes from the medical literature, and refers to analysis based on the original random assignment of individuals to treatment or control groups, regardless of whether or not individuals actually received or complied with the treatment. In our context, it gives the impact of assignment to migration status through the lottery, regardless of whether individuals who win the lottery actually migrate or not.

height as a pre-existing measure of health, and whether or not the applicant was born on the main island of Tongatapu, as a measure of having more urban skills. The addition of these controls reduces the size of the estimated effect only slightly, to $90, which is not significantly different from that obtained without controls. Column 3 controls further for past income, which is expected to also capture the effect of a host of unobserved individual attributes that determine income. The addition of this term only marginally changes the estimated intent-to-treat effect, which is now estimated to be $87. The fact that the estimated program effect changes only slightly in magnitude as we add the controls is consistent with the result in Table 1, which showed that the lottery succeeded in randomizing these controls across applicants and non-applicants.

3.3 Average treatment effects

These unbiased estimates of the ITT are substantially smaller than the biased estimate of the SEE-TT both because many individuals in the treatment group actually fail to receive the treatment (eg. migrate) and because of the potential dropout bias arising from non-random migration among those who do win the lottery. Heckman et. al. (2000) demonstrate that under the following assumptions: 1) lottery losers do not substitute for the migration treatment, 2) dropouts among the lottery winners are unaffected by winning the lottery, and 3) dropouts among the lottery winners have the same mean outcome as lottery losers who would have been dropouts if they had won the lottery; an unbiased estimate of the average treatment effect on the treated can be calculated which is adjusted for dropout bias (ADJ-TT):

$$\text{ADJ-TT} = \text{ITT} / p \qquad (7)$$

where p is the proportion of lottery winners who migrate. (eg the proportion of non-dropouts). Using the ITT of \$90.63 from column 1 in Table 2 and p=0.33 we can calculate that migrating increased the weekly work income of Tongans by \$274.

Instrumental variables provide another approach for estimating average treatment effects with experimental data Returning to equation (2), we can consistently estimate λ if an excluded instrument exists which is correlated with whether an individual migrates, $Migrate_i$, and is uncorrelated with the error term in this equation, ε_i. In our application, the PAC lottery outcome can be used as an excluded instrument because randomization ensures that success in the lottery is uncorrelated with unobserved individual attributes which might also affect income and success in the lottery is strongly correlated with migration (the first stage F-statistic is 61.5).[8] This estimate of λ is called the local average treatment effect (IV-LATE) and can be interpreted as the effect of treatment on individuals whose treatment status is changed by the instrument. In our application, this is the effect of migration on the income of individuals who migrate after winning the lottery. Angrist (2004) also demonstrates that in situations where no individuals who are assigned to the control group receive the treatment (eg. there is no substitution) then the IV-LATE is the same as the average treatment effect on the treated (IV-TT). In models with no covariates this also equals the ADJ-TT (Angrist, Imbens and Rubin, 1996).

[8] Validity of the instrument also requires that the lottery outcome does not directly affect incomes conditional on migration status. One could conceive of stories such as that winning the lottery and not being able to migrate causing frustration which leads individuals to work less, or conversely, winning the lottery acts as a spur to work harder in order to afford the costs of trying to find a job in New Zealand. However, such possibilities were not encountered in our field work, and as is seen in Table 2, income of non-migrants among the successful ballots is very similar to income of the unsuccessful ballots. This gives us reason to believe the instrument is a valid one.

Column 4 of Table 3 reports the IV-TT estimator when no other controls are included in the regression model, and estimates a gain in weekly work income of almost $274 from migrating, which is identical to the estimate above based on the ITT/p formula. Column 5 then adds the same control variables used above when estimating the ITT; the estimate increases slightly to $281. Column 6 adds past income as a further control, measured here as self-reported income from 2003. Past income is likely to capture a host of unobserved attributes of individuals which affect labor market performance and the likelihood of migrating conditional on winning the lottery, and is seen to be strongly significant. Each additional dollar of past income in 2003 is associated with 66 cents higher wage income today. Adding past income as a control results in an estimated income gain from migration of $274 per week. This is the same as was obtained in the model with no covariates, and confirms that randomization succeeded in making ballot success orthogonal to the other variables.

Therefore, after controlling for observable differences remaining after randomization, we estimate that a successful ballot increases expected income of PAC applicants by $91 per week, while migrating increases mean income by $274. Given that mean income of applicants with unsuccessful ballots is $104, this represents a 88% increase in expected income from winning the lottery, and a 263% increase in income from migrating.

4. Non-experimental estimators

The natural experiment provided by the use of a lottery to admit Pacific Islanders to New Zealand provides a unique opportunity to estimate the gain in income from migration.

Other studies of migration are forced to use non-experimental methods to attempt to deal with the selectivity issues associated with migration, comparing the incomes of migrants to that of non-migrants of similar observable characteristics. In this section we explore how well such methods work in practice, comparing the results obtained from different non-experimental methods to the experimental results described above.

This approach for studying the validity of non-experimental methods has a long history in the labor program evaluation literature. For example, in perhaps the first attempt to do so, Lalonde (1986) compared experimental estimates from the National Supported Work (NSW) Demonstration to non-experimental results calculated using control groups created from household survey data. For this program and treatment, Lalonde found that non-experimental methods did a poor job of replicating the experimental results. Heckman, Ichimura and Todd (1997), Dehejia and Wahba (2002), and Smith and Todd (2005) each further exploit the data collected for the NSW to examine whether particular refinements to non-experimental methods can lead to a better replication of the experimental results.

In summary, these papers demonstrate that more accurate non-experimental estimates can be achieved if the treatment and non-experimental control groups are: i) compared over a common support (eg. the distribution of the likelihood of receiving the treatment is similar in both groups), ii) located in the same labour markets, and iii) administered the same questionnaire (eg. data is collected from both groups in an identical manner). A significant improvement can further be achieved if data is collected from both the preand post-treatment periods and a ‘difference-in-differences’ estimator is used to control for unobserved differences between the treatment and control groups by differencing out individual fixed effects which are correlated with both the outcome and the likelihood of being treated. Nonetheless, even with these refinements, Smith and Todd (2005, p.305) conclude, “Our analysis demonstrates that while propensity score matching is a potentially useful econometric tool, it does not represent a general solution to the evaluation problem.”

Recall that PINZMS collects data for a sample of non-applicants to the lottery selected from either the same villages that the migrants had been living in prior to migrating or in the same villages that unsuccessful ballots were found in and administers them an identical questionnaire to the one given to other non-migrants in our sample (eg. the experimental control group). Thus, these individuals serve as a perfect non-experimental control group on which to test alternative methodologies for estimating the gains from migration. As discussed above, all individuals in our sample report their income from the previous year allowing us to also implement a ‘difference-in-differences’ estimator.

Before proceeding with microeconomic non-experimental estimators, it is worth comparing the experimental estimate of the income gains to the cross-country macro estimator. Cross-country studies of the determinants of migration often use differences in per capita national income as proxies for the income gains from migration (e.g. Clark, Hatton and Williamson, 2002). In 2004, New Zealand’s GDP per capita was NZ$30,469, while Tonga's was NZ$2,044.[9] This difference in GDP per capita therefore equates to NZ$546 per week, or twice as large as the actual gain experienced by the average migrant in our survey.

4.1. The Single Difference Estimator

We begin by examining whether a simple single difference estimate calculated using only information from the migrant group provides a good estimate of the income gains from migration. Several recent surveys of new immigrants (eg. the Longitudinal Immigrant Survey: New Zealand (LisNZ); and the New Immigrant Survey (NIS) in the U.S.) ask about income prior to migration. Thus, one approach to estimating the average income gain from migration is to calculate the mean difference between the migrant's pre-migration and post-migration incomes. That is, the estimate is:

$$\lambda_{SD} = E[\text{Income}_{i,t} - \text{Income}_{i,t-1} \mid i \text{ migrating between } t \text{ and } t{-}1] \quad (8)$$

Adding time subscripts and control variables to equation (2), and assuming that slope coefficients do not change over time, we have:

$$\text{Income}_{i,t} = \mu_t + \lambda * \text{Migrate}_{i,t} + \pi' X_{i,t} + \eta_{i,t} \quad (9)$$

Then we see that:

$$\lambda_{SD} = (\mu_t - \mu_{t-1}) + \lambda + E[\pi'(X_{i,t} - X_{i,t-1})] \quad (10)$$

There are two possible sources of bias in such an estimate. Firstly, if individuals on average change their attributes, such as experience, or education, then we would expect

[9] Source: World Bank GDF and WDI Central (August 2005 update) for population and GDP. The exchange rate of 1 pa'anga to 0.729 NZ$ prevailing at the time of our survey was used to convert Tongan GDP per capita to New Zealand dollars.

their incomes to change over time and so the third term to be non-zero. Secondly, if there are overall macroeconomic movements, mean income for those not migrating will differ from one period to the next. This re-emphasizes the fact that the counterfactual one would ideally like is what a given individual would be earning in the current time period if he or she didn't migrate; this could be different from what they earned before migration due to macroeconomic factors or changes in the income-earning potential of the individual over time. A third potential form of bias when it comes to estimation is that previous income is likely to be subject to greater recall error than current income.

The first row of Table 4 provides the estimate λ_{SD}, calculated as the difference between the current income of our migrant sample and what they reported earning prior to migration. Based on this method, we would estimate an income gain of \$341. Comparing this to columns 4 and 6 of Table 3, we calculate that this method results in estimated income gains which are 25% higher than the experimental estimate.

We can examine the magnitude of the first source of bias in this estimator by examining the increase in income that occurred for the unsuccessful ballots, who remained in Tonga. Mean income increased \$28 per week for this group, which accounts for 42 percent of the difference in income gains estimated via this method compared to the experimental estimates.

4.2. OLS

A second non-experimental method commonly used to estimate the returns from migration is to assume that all differences between migrants and non-migrants which affect income are captured by the regressors in an OLS regression. One then estimates λ through the following regression:

$$Income_i = \kappa + \lambda^*Migrate_i + \pi' X_i + \upsilon_i \quad (11)$$

We estimate equation (11) by combining the sample of migrants in New Zealand with the sample of non-applicants in Tonga. We do this for two samples in Tonga. One individual from each household of non-applicants was asked a longer set of questions, including information on their family networks in New Zealand, expectations about the future, and other broader issues. This was done for the group of pseudo-applicants, consisting of the oldest member aged 25 to 35 in the non-applicant household (or oldest member aged 18-45 if the household did not have a 25 to 35 year old). The first sample we use combines these individuals with the migrants. The second sample uses all individuals aged 18 to 45 in the non-applicant households. The set of controls used in equation (11) are the same as used above, and include age, education, marital status, sex, birthplace and height.

Table 4 shows that this results in an estimated income gain from migration of $384 using the restricted sample, and an income gain of $360 using the wider sample. Appendix 1 provides the full regression results. Comparing these with the experimental estimates, we see that the restricted sample overestimates the income gain by 40% and the full sample overestimates the income gain by 31%. The direction of this bias is consistent with the view that migrants have more drive or greater labor market ability than non-migrants.

Column 2 of Appendix 1 repeats this regression for the full-sample of 18 to 45 year olds without including any of the X variables as controls in equation (11). The coefficient on migration is \$386. Adding the observable characteristics as controls in column 3 reduces this to \$360, showing positive selection on observables. However, the change in the migration coefficient from adding these controls is not significant, and their addition only reduces the overestimation of the income gains from 41% to 31%. It therefore seems that most of the OLS bias is due to selection on unobserved characteristics.

4.3. Difference-in-Differences

Using self-reported past income, we can also control for time invariant individual attributes which affect labor market income via difference-in-differences regression. Since we do not have panel data on all the control variables, we estimate the following version of the difference-in-differences regression :

$$Income_i - PastIncome_i = \kappa + \lambda * Migrate_i + \pi' X_i + \upsilon_i \qquad (12)$$

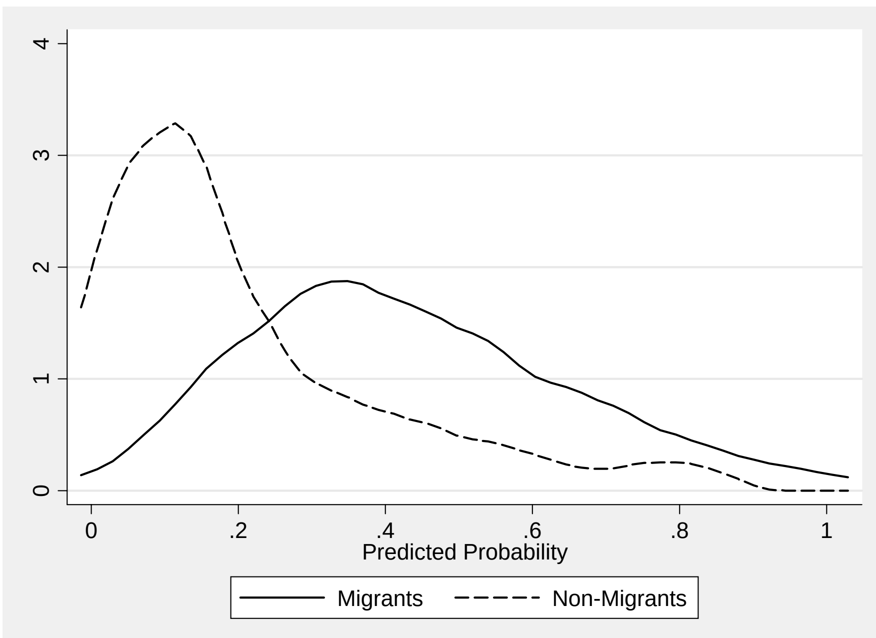
Controlling for past income lowers the estimated income gain to \$375 using the restricted sample and \$328 using the wider sample. Columns 4 and 5 of Appendix 1 provide the full set of coefficients. These estimates are now respectively 37% and 20% higher than the experimental estimate, although given our sample sizes, we can only reject equality with the experimental estimate for the narrower sample. There are two main possible sources of remaining bias. The first is that unobserved characteristics like drive and ability may be rewarded differently in the New Zealand and Tongan labor markets, so that individual effects are time-varying. The second is that we are comparing migrants to not-very-similar non-migrants, and so the assumption of a common underlying trend in labor income is not tenable. The latter assumption is eased by using the wider sample, and can be relaxed further by ensuring that the migrants are compared to sufficiently similar non-migrants, which the next method attempts to do.

4.4. Propensity-Score Matching

Propensity-score matching is perhaps the non-experimental evaluation technique which has attracted most research interest in recent years, with proponents claiming that it can replicate experimental benchmarks when appropriately used (Dehejia and Wahba, 2002; Dehejia 2005). Estimation takes place by first estimating a probit equation for the probability of migrating, and then matching each migrant to non-applicants with similar predicted probabilities of migration. This enables migrants to be compared to individuals who are similar in terms of observed characteristics. Once the matches are constructed, the gain in income is calculated as the mean income for migrants less the mean income for the matched sample. We use the nearest-neighbor matching, and following Abadie et al. (2001) match each migrant to the four nearest neighbors.

Our base specification uses the same set of control variables as used in the regression analysis to form the match. The existing literature (Heckman, Ichimura and Todd, 1997; Smith and Todd, 2005) have noted that difference-in-difference matching estimators can perform substantially better than cross-sectional matches. While we do not have panel data on all matching variables, the inclusion of past income allows us to obtain estimates similar in spirit to difference-in-difference matching. Figure 1 then shows kernel densities of the propensity scores when past income is included alongside the other regression controls in forming the match. Note that there is considerable overlap in the distributions, with some migrants and some non-applicants in almost all the range. The propensity score for the migrant group ranges from 0.069 to 0.947, while that of the non-applicant comparison group ranges from 0.000 to 0.789. Estimation is restricted to the area of common support, where the two distributions overlap.

Figure 1: Propensity Scores for Migrants and Non-migrants



One potential criticism is that these base specifications are relatively parsimonious, using only 6 or 7 covariates to form the match. This is in large part due to the need to use retrospective questions and time invariant attributes to form the match, since the survey was cross-sectional. To investigate the robustness of the matching results to a more flexible specification, we also estimated the propensity score allowing for interactions of sex with each of the other covariates, quartics in age and years of schooling, and an interaction between age and education, for a total of 19 covariates.

For each of these three specifications of variables used to form the match we calculate the sample average treatment effect (SATE) and sample average treatment effect for the treated (SATT) following Imbens (2004). Table 5 reports these estimates in rows A, B and C. [10] Once we control for past income, the SATE and SATT are very similar to one another. We focus on the SATT, since this is more directly comparable to the experimental treatment effect estimated using the migration lottery as an instrument for migration.

Under the basic specification, the estimated income gain is $364 per week, 33% higher and significantly different from the experimental estimate of $274. Adding past income as a control lowers the bias to 28.5% and adding interactions reduces it to 27.4%. The t-statistic for testing equality of the treatment effect in model C with the experimental estimate is 1.61, close to the margin of being able to reject equality at the 10% level of significance.

However, Abadie and Imbens (2005) show that matching estimators are in general not root-N consistent when more than two continuous covariates are used for the matching. Intuitively as more covariates are used, it becomes more difficult to obtain a close match to the treated observation. This results in a conditional bias term of stochastic order $N^{-1/k}$, where k is the number of continuous matching variables. They propose a bias-corrected

[10] Propensity-score matching was estimated in STATA using the routine described in Abadie et al. (2001).

matching estimator which corrects for this and is root-N consistent. Our base model in A has 3 continuous covariates (age, education and height), while the model with interactions in C has 14 continuous covariates. We therefore implement the bias-correction and report the bias-adjusted SATT in Table 5.

The bias-adjustment brings the matching treatment effects closer to the experimental estimate, and we can no longer reject equality. In model C, with interactions, the bias is reduced from 27.4% to 19.9%. Dehejia (2005) notes that sensitivity of the matching estimator to small changes in the specification used is one diagnostic as to the quality of the comparison group. The bias-adjusted estimators are not that sensitive to the particular specification used for matching, ranging from $329 to $346 per week as the estimated income gain. Based on this, one would therefore be likely to conclude that the matching technique is working reasonably well in this context, even without reference to the experimental data.

Rows D and E of Table 5 conduct two other robustness tests suggested by the literature. The first is to not only estimate the matching estimator over the area of common support, but also to examine robustness to trimming observations in the support with very low or very high probabilities of being selected. Panel D trims propensity scores which are less than 0.01, 0.05, 0.10 and 0.15 or greater than 0.99, 0.95, 0.90 and 0.85 respectively. After the bias-adjustment, the estimated treatment effect is not very sensitive to such trimming, resulting in a bias of 18.9 to 20.1 percent. The second robustness test examines the sensitivity of the estimator to the number of neighbors used in forming the match. This trades bias for efficiency, which is seen in the smaller standard errors when more neighbors are used. Again the point estimates are very robust to this choice of specification, and result in a 20% higher income gain than is estimated by the experiment.

Is there a pre-migration-lottery earnings dip?

In studies of labor training programs, Heckman, Ichimura and Todd (1997) and Dehejia and Wahba (2002) note the importance of including information on labor force histories in estimating the probability of participation when using matching estimators. A particular concern in evaluating labor training programs is the dip in earnings often observed prior to participation in such programs (Ashenfelter, 1978). For this reason, Dehejia (2005) stresses that two or more years of pre-treatment earnings are desirable for use in matching. We only measure income for one period prior to migration for the migrants in our sample, and so are unable to use two or more periods of pre-lottery earnings in our matching. We do, however, have several years of labor histories for the sample observed in Tonga, and so can investigate whether there is a dip in earnings prior to applying for the migration lottery by looking at earnings of lottery losers.

After deflating by the Tongan Consumer Price Index to convert earnings into March 2002 pa'anga, we find that applicants to the lottery were experiencing very moderate income *growth* in the run-up to their application for the lottery. Mean real weekly income was T78.78 in 2002, T81.98 in 2003 and T84.46 in 2004. This may still represent an earnings dip compared to the counterfactual of not-applying for the lottery if the economy as a whole is growing. Therefore, to rule this out, we match lottery losers to non-applicants on the basis of the same set of basic controls used in specification A in Table 5, along with real weekly work income in 2002 and 2003. We then can ask whether individuals who would apply for the lottery in early 2005 had lower income in 2004 than similar individuals, with similar incomes in 2002 and 2003, who did not apply for the lottery. The estimated mean difference in weekly income in 2004 from this match is –1.63 pa'anga, with a standard error of 11.35. This is statistically insignificant and amounts to less than 2 percent of mean weekly income in 2004 of lottery applicants. Therefore it appears that we can rule out a pre-migration lottery dip in earnings. Hence the matching estimators presented in Table 5 should not be greatly affected by the use of one rather than two year's pre-treatment earnings.

4.5. Instrumental Variables with a Non-experimental instrument

Like the regression approach, propensity score matching relies on selection on observables, so will overstate the income gains if migrants are more talented or have more drive than observationally similar non-migrants. An alternative approach to non-experimental estimation of the impact of migration explicitly recognizes that migrants are likely to be non-randomly selected, even conditioning on observables, and so attempts to find instruments for migration. An example is Munshi (2003), who uses rainfall in Mexican villages as an instrument for migration when looking at the effect of migration networks on job outcomes in the United States. Given the small size of Tonga, weather variation does not provide an instrument in our application. We instead consider two potential instruments, with varying likelihoods of the exclusion restriction being satisfied.

Several studies looking at the impact of migration on the sending country have employed historic migration networks (e.g. Woodruff and Zenteno 2001, McKenzie and Rapoport 2004). In our context it is likely that having a large network of relatives in New Zealand helps predict whether an individual migrates, and so we consider the effect of using the total number of types of relatives an individual has in New Zealand as an instrument for migration. This is strongly correlated with migration (first-stage F-statistic is 14.3). However, we would be highly concerned that the exclusion restriction is violated for this instrument, since many individuals in our survey said they found their first job in New Zealand through relatives.

We investigate this by estimating equation (11) using the migrant network as an instrument for migration, restricting analysis to the sub-sample for which we have information on their network. Table 4 and column 6 of Appendix 1 show that this results in an estimate of the income gain from migration of $499, which is 82% higher than the experimental estimate. Thus using a poor instrument results in a bias almost as large as the cross-country estimator.

A valid instrument is a variable which predicts whether or not people apply to migrate, but doesn't otherwise affect their labor market outcomes if they move to New Zealand. Our survey asked eligible individuals who didn't apply for the Pacific Access Category why they didn't apply. The most important reason given for not applying was that they did not know the requirements, which 98% of non-applicants listed as a very important reason for not applying. A distant second among the other reasons given was that they didn't think the chances of getting selected in the lottery were very high, which 12% listed as very important, and a further 60% gave as somewhat important.

This motivates our choice of an alternative potential instrument, which is based on how close the individual's house in Tonga is to the New Zealand Immigration Service (NZIS) office. Information about the requirements of the Pacific Access Category is obtained from this office, and the applications have to be delivered there. GPS coordinates were taken of each of the households in our survey, and of the NZIS office location, and based on these, the (log) of the distance between each household and the NZIS office was calculated.[11]

Comparing migrants to non-applicants, we find log distance to be a very strong predictor of migration, with a first stage F-statistic of 21.9. Row 5 in Table 4 and column 7 of Appendix 1 show the resulting estimate of the income gains from migration using log distance as an instrument. The estimated gain is $305, which is 11% higher than the experimental estimate. Since migration status is a binary variable, an alternative approach to the two-stage least squares (2SLS) model is to use full information maximum-likelihood (FIML) to estimate a treatment effects model with an endogenous binary

[11] A possible threat to the exclusion restriction needed for distance to be a valid instrument is that individuals who lived further away from the NZIS office would have lived in more isolated, less urban areas. If this makes them less able to adapt to city life in New Zealand, then we would expect an upward bias in the IV estimator, since it would be in part capturing the returns to more urban experience. To investigate this possibility, for the group of migrants in New Zealand we regressed income on our set of controls, including past income, and log distance. The coefficient on log distance is positive (9.0) and insignificant (p-value of 0.76). Thus there is no strong effect of living in a location closer to the NZIS office in Tonga on labor income, and if anything, migrants that lived further away earned slightly more in New Zealand. Based on this it is likely that one would conclude that this was a reasonable instrument, even without reference to the experimental comparison.

treatment.[12] Column 8 in Appendix 1 shows the models coefficients are similar to those from 2SLS. Row 6 of Table 4 shows a slight reduction in the estimated gain from migration, down to \$298, which is only 9% higher than the experimental estimate.

5. Exploring Selection Directly

The non-experimental methods all overestimate the income gains from migration compared to the estimate obtained from the lottery. This suggests that lottery applicants are positively selected in terms of unobserved individual characteristics which determine labor market earnings. Selection could occur both in terms of the decision to apply for the lottery, and the decision to migrate conditional on winning the lottery. We examine here the evidence for each type of selection.

We first examine the overall extent of selection by comparing the pre-migration income of migrants to that of observationally similar non-applicants via the following regression:

$$Income_{i,t-1} = \alpha + \beta*Migrant_{i,t} + \gamma'X_{i,t-1} + \varepsilon_{i,t-1} \quad (13)$$

where *X* consists of a set of controls, such as age, education, gender, marital status, height, and migrant network, and *Migrant* is a dummy variable taking the value one if person *i* applies for the lottery and migrates in the next period, and zero if they don't apply for the lottery.

We then consider selection into the lottery by estimating this for the comparison of all lottery applicants to all non-applicants, replacing the *Migrant* dummy variable with a dummy variable for applying for the lottery. We compare the income for migrants in the

[12] This was carried out using the treatreg command in STATA.

12 months prior to migration to the income of non-applicants in 2003, which corresponds to a similar reference period. The coefficient β then indicates whether migrants or applicants earned more or less prior to applying for the lottery than non-applicants, conditional on their observed characteristics. As with the non-experimental estimators, we carry out this analysis for the two groups of non-applicants: all individuals aged 18-45, and the set of pseudo-applicants.

The first two columns of Table 6 report the results of estimating equation (13), comparing migrants to all 18-45 year old non-applicants. The coefficient β is positive and highly significant. Migrants and non-applicants are seen to differ both in terms of observables and unobservables. Controlling for observables lowers the difference in lagged income from $56 per week to $29 per week. However, given that the average income of non-applicants in this group is $33 per week, we see that migrants earned almost twice as much as observationally similar non-applicants prior to them migrating. Similar results are shown in columns (3) and (4), where we consider selection into the lottery and compare all principal applicants to non-applicants. We can not reject equality of the coefficient on migrating in column (2) with the coefficient on applying in column (4).

In Columns (5) and (6) we compare migrants to the pseudo-applicant group. Despite the smaller sample, we still find a statistically significant positive coefficient on the migration dummy. The average income for the pseudo-applicants was $61 per week, so migrants are estimated to have earned over 35% more than observationally equivalent non-applicants in the pseudo-applicant group.

In the last two columns of Table 5 we modify equation (13) to examine whether there is selective compliance to the treatment of winning the lottery, comparing the pre-migration incomes of lottery winners who migrate to lottery winners who had not migrated at the time of the survey. The coefficient on migrating is found to be very close to zero in magnitude ($13 per week without controls and $7 a week with controls), and insignificant, with t-statistics below 0.9 in absolute value. Migrants therefore do not appear to differ greatly from non-migrant lottery winners in terms of unobservable characteristics affecting labor market earnings.

These results therefore suggest that lottery applicants are positively selected in terms of unobservable characteristics which affect labor market earnings, but that there is little evidence for selective compliance to the treatment of winning the lottery, conditional on applying. This positive selection concurs with the differences seen in the comparison of the experimental and non-experimental estimates of the income gains from migration.

6. Cost-of-living Adjusted Income Gains

The experimental estimate of an average gain in income of $274 per week from migrating is based on comparison of incomes earned in New Zealand to incomes earned in Tonga, converted into New Zealand dollars using the exchange rate prevailing at the time of the survey, of 1.372 Pa’anga per New Zealand Dollar (P/NZD). To investigate the sensitivity of the size of the income gain to adjustments for differences in the cost of living, we collected price data in both countries and formed several purchasing power parity exchange rates.

The first measure, inspired by the Big Mac index of *The Economist*, is the PPP exchange rate based on the price of KFC takeaway chicken in New Zealand and a similar product in Tonga.[13] This gives a rate of 1.669 P/ NZD. The second measure is based on food prices for the main foods consumed in our surveys, using the weights from the Tongan Consumer Price Index. This gives a rate of 1.050 P/NZD, reflecting the cheaper cost of some foods, such as fish, taro, and coconut in Tonga. The third measure uses food, transport and durable goods, giving a rate of 1.57 P/NZD, reflecting that most fuel and durable goods are imported into Tonga and are more expensive there. Finally, our preferred measure adds an allowance for housing services to the third basket, to arrive at a basket close to a full consumer price index. This gives a PPP exchange rate of 1.368 P/NZD, which is remarkably close to the prevailing exchange rate.

Therefore, based on these measures, one estimates that the experimental estimate of the income gains from migration in terms of PPP-adjusted New Zealand dollars range from \$227.7 using the food PPP rate, to \$300.44 using the KFC index. Using the preferred PPP rate the estimate is \$273.3, which is very close to that attained using the market exchange rate. Hence our estimated gain seems very robust to cost-of-living adjustments.

[13] There is no McDonalds in Tonga, precluding the use of a Big Mac index. However, KFC is very popular amongst Tongans in New Zealand, and a close copy of KFC called "Country Fried Chicken" operates in Tonga with similar products.

7. Conclusions

The lottery used to select migrant applicants to New Zealand from the Pacific Islands provides a unique natural experiment which can be exploited to estimate the income gains from migration and to examine how successful are non-experimental methods in estimating these gains. We estimate that winning the lottery increases expected weekly income by NZ $87, while migrating increases mean income by $274. Migration therefore results in an increase in work income of over 263%.

Our results show selection to be important in measuring the income gains from migrating, with migrants positively selected in terms of both observable characteristics and unobserved labor market attributes such as ability and drive. Direct evidence for this is observed from comparisons of prior income between migrants and non-applicants. As a result of this selection, we find that non-experimental estimation methods all overstate the income gains from migration, by between 9 and 82 percent. Among the non-experimental methods, we find that a good instrument (log distance to the office where ballots are deposited) works well, while a poor instrument (migrant network) works very badly. Difference-in-differences and propensity score matching with bias-adjustment work best of the non-IV non-experimental methods, but both overestimate the income gains from migration by around 20 percent. Nevertheless, these methods do work better than OLS, and are not statistically different from the experimental point estimate with our sample sizes.

References:

Abadie, A., D. Drukker, J.L. Herr and G.W. Imbens (2001) "Implementing matching estimators for average treatment effects in Stata", *The Stata Journal* 1(1): 1-18.

Abadie, A. and G.W. Imbens (2005) "Large sample properties of matching estimators for average treatment effects", forthcoming, *Econometrica.*

Angrist, J.D. (2004) "Treatment Effect Heterogeneity In Theory and Practice", *Economic Journal* 502: C52-C83.

Angrist, J.D., G.W. Imbens and D.B. Rubin (1996) "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association* 91(434): 444-55.

Ashenfelter, O. (1978) "Estimating the effects of training programs on earnings", *Review of Economics and Statistics* 60: 47-57.

Borjas, George J. (1987) "Self-selection and the earnings of immigrants", *American Economic Review* 77(4): 531-53.

Chiquiar, D. and G. Hanson (2005) "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States", *Journal of Political Economy* 113(2): 239-81.

Clark, X., T.J. Hatton and J.G. Williamson (2002) "Where do U.S. Immigrants come from and why?", *NBER Working Paper 8998.*

Deaton, A. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Johns Hopkins University Press, Washington DC.

Dehejia, R. (2005) "Practical propensity score matching: a reply to Smith and Todd", *Journal of Econometrics* 125(1-2): 355-64.

Dehejia, R. and Wahba, S. (2002) "Propensity Score Matching Methods for Non-Experimental Causal Studies", *Review of Economics and Statistics* 84(1): 151-161.

Hartog, J. and Winkelmann, R. (2003) "Comparing Migrants to Non-migrants: The Case of Dutch Migration to New Zealand", *Journal of Population Economics* 16(4): 683-705.

Heckman, J., Hohmann, N., Smith, J. and Khoo, M. (2000) "Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment", *Quarterly Journal of Economics* 115(2): 651-694.

Heckman J., Ichimura H. and Todd P. (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economic Studies* 64 (4): 605-654.

Imbens, G. (2004) "Nonparametric estimation of average treatment effects under exogeneity: a review", *Review of Economics and Statistics* 86(1): 4-29.

Lalonde R (1986). "Evaluating the Econometric Evaluations of Training Programs." *American Economic Review* 76: 604–620.

LaLonde, R. and Topel, R.. (1997) "Economic Impact Of International Migration And The Economic Performance Of Migrants", In M. Rosenzweig and O. Stark (ed.) *Handbook of Population and Family Economics*, Elsevier Science, pp. 799-850.

McKenzie, D., and Rapoport, H. (2004) "Network Effects And The Dynamics Of Migration And Inequality: Theory And Evidence From Mexico", *Stanford University Center for International Development (SCID) Working Paper 191*.

Munshi, K. (2003) "Networks in the Modern Economy: Mexican Migrants in the United States Labor Market", *Quarterly Journal of Economics* 118(2): 549-597.

Robinson, C. and Tomes, N. (1982) "Self-selection and Inter-provincial Migration in Canada", *Canadian Journal of Economics* 15(3): 474-502.

Rosenbaum, P. and Rubin, D. (1983) "The central role of the propensity score in observational studies for causal effects", *Biometrika* 70: 41-55.

Smith J and Todd P (2005). "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators." *Journal of Econometrics* 125 (1-2): 305-353.

Woodruff, C. and Zenteno, R. (2001). "Remittances and Microenterprises in Mexico." Working Paper, UCSD and ITESM-Guadelejara, December.

TABLE 1: TEST FOR RANDOMIZATION

Comparison of Ex-ante characteristics of principal applicants in successful and unsuccessful ballots

	Sample Means APPLICANTS Successful Ballots	Sample Means APPLICANTS Unsuccessful Ballots	T-test of equality of means p-value
Age	33.6	33.7	0.91
Years of schooling	11.9	11.5	0.37
Proportion male	0.55	0.51	0.52
Proportion born on Tongatapu	0.75	0.79	0.54
Proportion who had been to NZ before 2000	0.39	0.35	0.63
Proportion who are married	0.60	0.62	0.77
Height	171.6	169.3	0.16
Proportion selling fish in 2003	0.03	0.06	0.40
Proportion selling crops in 2003	0.22	0.26	0.52
Income in 2003/before moving	103.7	88.0	0.32
Proportion with the following family members living in NZ at time of last application:			
Father/Father-in-law	0.38	0.44	0.45
Mother/Mother-in-law	0.40	0.35	0.46
Brother/Brother-in-law	0.72	0.71	0.78
Sister/Sister-in-law	0.64	0.60	0.63
Aunt or Uncle	0.65	0.55	0.17
Total Sample Size	**120**	**78**	

Average first quarter 2005 exchange rate of 1 Pa'anga = 0.729 NZ Dollars used for comparing mean incomes

TABLE 2: SAMPLE MEANS OF EMPLOYMENT, HOURS WORKED AND WAGES

	Observations	Proportion Employed	Mean hours worked per week	Mean weekly income from work (NZ Dollars)
APPLICANTS	198	0.723	27.3	108.9
Successful Ballots	120	0.662	28.4	194.7
Migrants	65	0.754	33.3	424.5
Non-migrants	55	0.618	26.0	81.1
Unsuccessful Ballots	78	0.731	27.1	104.1
NON-APPLICANTS	60	0.672	24.2	69.5
All Non-applicants 18-45	180	0.439	16.2	41.4
T-tests of equality of means				
Successful Ballots vs Unsuccessful Ballots		0.349	0.683	0.000
Migrants vs Non-migrant Successful Ballots		0.111	0.086	0.000
Migrants vs unsuccessful ballots		0.754	0.105	0.000
Pure Experimental Estimators of the Gain in Income from Migration				
Intention-to-treat effect			90.6	
SEE-TT			320.4	

Notes:
Average first quarter 2005 exchange rate of 1 Pa'anga = 0.729 NZ Dollars used for comparing mean incomes
SEE-TT is the simple experimental estimator of the effect of the treatment on the treated, and compares migrants to unsuccessful ballots.

TABLE 3: REGRESSION-BASED EXPERIMENTAL ESTIMATES

Dependent Variable: Weekly Income from Work in New Zealand Dollars

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Ballot Success Dummy	90.634	89.741	87.390			
	(3.68)**	(3.71)**	(3.89)**			
Male Dummy		-29.070	-23.855		-33.104	-27.772
		(1.19)	(1.08)		(1.43)	(1.33)
Married Dummy		-4.493	24.535		-10.695	18.376
		(0.16)	(1.05)		(0.40)	(0.82)
Age Dummy		0.558	-0.886		0.987	-0.462
		(0.34)	(0.71)		(0.64)	(0.41)
Years of Education		13.427	4.605		12.034	3.274
		(2.03)*	(1.18)		(1.99)*	(0.91)
Born on Tongatapu Dummy		29.167	27.600		29.594	28.005
		(1.55)	(1.87)		(1.64)	(2.04)*
Height		1.281	0.381		1.249	0.353
		(1.96)	(0.92)		(2.04)*	(0.93)
Past income			0.662			0.660
			(6.98)**			(7.31)**
Migration Dummy				273.996	281.050	273.736
				(4.46)**	(4.56)**	(4.99)**
Constant	104.051	-297.878	-60.422	104.051	-285.011	-48.595
	(8.85)**	(2.45)*	(0.74)	(8.90)**	(2.45)*	(0.66)
First stage F-statistic on instrument:				66.53	61.88	61.51
Observations	197	191	190	197	191	190
R-squared	0.04	0.14	0.27			

Robust t statistics in parentheses

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

Average first quarter 2005 exchange rate of 1 Pa'anga = 0.729 NZ Dollars used for comparing mean incomes

TABLE 4: NON-EXPERIMENTAL ESTIMATES

Method:	Estimate	s.e.	Percent difference compared to experimental estimate 273.996	Testing equality with 273.996 T-stat
1) Using pre-migration income as the counterfactual	341.3	46.4	24.6	1.45
2) Selection on Observables: OLS regression				
oldest member aged 18-45	383.5	46.4	40.0	2.36
all members aged 18-45	360.0	41.2	31.4	2.09
3) Difference-in-Difference Regression				
oldest member aged 18-45	375.2	46.4	36.9	2.18
all members aged 18-45	328.5	42.8	19.9	1.27
4) Instrumental Variables using migrant network	498.8	209.6	82.0	1.07
5) Instrumental Variables using log distance to NZIS office	305.0	93.5	11.3	0.33
6) FIML treatment effects model using log distance to NZIS office as an instrument	298.0	69.3	8.8	0.35

Notes: Experimental estimate is the IV estimate from column 6, Table 3.

TABLE 5: PROPENSITY-SCORE MATCHING ESTIMATES

	Estimate	s.e.	Percent difference compared to experimental estimate 273.996	Testing equality with 273.996 T-stat
A: Matching without using past income				
SATE	335.7	40.2	22.5	1.53
SATT	364.0	44.0	32.9	2.04
bias-adjusted SATT	346.3	45.4	26.4	1.59
B: Matching using past income				
SATE	355.8	43.6	29.9	1.88
SATT	352.2	45.4	28.5	1.72
bias-adjusted SATT	333.4	46.2	21.7	1.29
C: Matching using past income and interactions				
SATE	346.2	44.7	26.3	1.62
SATT	349.1	46.5	27.4	1.61
bias-adjusted SATT	328.6	47.3	19.9	1.16
D: Trimmed, bias-adjusted SATT using specification C				
Trimming 0.01 and 0.99	328.9	47.1	20.0	1.17
Trimming 0.05 and 0.95	329.1	47.4	20.1	1.16
Trimming 0.10 and 0.90	328.5	48.3	19.9	1.13
Trimming 0.15 and 0.85	325.7	49.9	18.9	1.04
E. bias-adjusted SATT in specification C with different numbers of matches				
Nearest neighbor	330.1	59.0	20.5	0.95
Nearest 2 neighbors	330.6	50.9	20.7	1.11
Nearest 5 neighbors	328.6	45.8	19.9	1.19
Nearest 10 neighbors	330.4	43.0	20.6	1.31

Notes: A matches on gender, age, marital status, years of education, place of birth and height. B also includes past income, while C adds interactions of sex with each covariate, quartics in age and years of schooling, and an interaction between age and years of education. Estimation uses the 4 nearest neighbours for matching each observation, except in E.

Table 6: A Direct Look at Selection
Dependent Variable: Labour Income prior to applying/migrating

	Selection into the Lottery						Selection into Migration	
	All 18-45 year olds				Pseudo-applicants		Lottery Winners	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Migration Dummy Variable	55.95	29.46			28.28	21.83	13.14	7.43
	(5.30)***	(3.13)***			(2.16)**	(1.97)*	(0.86)	(0.51)
Applicant Dummy Variable			47.12	23.06				
			(6.53)***	(3.06)***				
Controls	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.145	0.330	0.099	0.205	0.037	0.331	0.006	0.129
Number of Observations	234	221	366	350	120	117	119	117

Notes:
Absolute value of t-statistic in parentheses, using heteroskedasticity-consistent standard errors
*, **, and *** indicate significance at the 10%, 5% and 1% levels
Controls include gender, age, marital status, years of education, place of birth and height. The total number of relative types in New Zealand is also used as a control in the pseudo-applicant and lottery winner regressions.

APPENDIX: NON-EXPERIMENTAL REGRESSIONS

Dependent Variable: Weekly work income for Columns 1-3 and 6-8.

Current Weekly work income - past weekly work income for columns 5 and 6.

	(1) OLS	(2) OLS	(3) OLS	(4) DD	(5) DD	(6) IV	(7) IV	(8) FIML
Dummy for Migration	383.490	385.880	360.009	375.226	328.498	498.797	304.999	298.034
	(8.27)**	(8.41)**	(8.73)**	(8.09)**	(7.68)**	(2.38)*	(3.26)**	(4.30)**
Male Dummy	176.214		75.998	162.739	80.517	200.683	97.477	97.608
	(2.62)*		(3.29)**	(2.29)*	(3.70)**	(2.36)*	(4.70)**	(4.01)**
Married Dummy	-115.798		-31.548	-125.159	-45.883	-109.672	-39.565	-39.703
	(1.87)		(1.15)	(1.94)	(1.77)	(1.84)	(1.60)	(1.52)
Age	4.706		2.797	2.314	0.083	4.807	0.872	0.824
	(1.38)		(1.57)	(0.67)	(0.04)	(1.44)	(0.49)	(0.45)
Years of Education	-2.056		-2.589	-17.712	-2.940	-10.519	2.531	2.717
	(0.18)		(1.03)	(1.44)	(1.30)	(0.60)	(0.82)	(0.63)
Born on Tongatapu	74.661		38.288	43.119	32.406	63.531	49.717	50.244
	(1.75)		(1.76)	(1.00)	(1.40)	(1.44)	(2.10)*	(1.92)
Height	7.094		3.589	7.404	5.964	6.397	5.960	5.981
	(2.38)*		(2.56)*	(2.42)*	(3.89)**	(2.07)*	(4.29)**	(4.26)**
Past Income						0.031	0.299	0.313
						(0.10)	(1.02)	(1.24)
Constant	-1,393.759	41.906	-664.329	-1,211.553	-1,022.750	-1,253.125	-1,478.666	-1,094.673
	(2.88)**	(9.94)**	(3.09)**	(2.36)*	(3.95)**	(2.47)*	(3.41)**	(4.22)**
Sample	oldest 25-35	all 18-45	all 18-45	oldest 25-35	all 18-45	oldest 25-35	all 18-45	all 18-45
Instrument						network	log distance	log distance
First stage F-statistic						14.24	21.93	
Number of Observations	118	230	230	116	226	116	226	226
R-squared	0.48	0.45	0.51	0.45	0.45			

Note: Oldest member 18-45 in the household is used when the household doesn't contain a 25-35 year old.

Robust t statistics in parentheses

Average first quarter 2005 exchange rate of 1 Pa'anga = 0.729 NZ Dollars used for comparing mean incomes

* and ** indicate significance at the 5% and 1% levels respectively.