Accounting for Selectivity and Duration-Dependent Heterogeneity When Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas

Author(s): John Gibson /David McKenzie /Steven Stillman

Source: *Economic Development and Cultural Change*, Vol. 61, No. 2 (January 2013), pp. 247-280
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/10.1086/668276
Accessed: 23/06/2013 16:49

Accounting for Selectivity and Duration-Dependent Heterogeneity When Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas

JOHN GIBSON
University of Waikato

DAVID MCKENZIE
Development Research Group, World Bank

STEVEN STILLMAN
University of Otago

I. Introduction

The impacts of international emigration and remittances on incomes and poverty in sending areas are increasingly studied with household survey data. Empirical analysis is needed because the effect of emigration is a priori unclear. Households with emigrants typically benefit from remittance inflows. There are also fewer mouths to feed among household members left behind. On the other hand, earnings and other household inputs that emigrants would have generated locally are lost. Since it is typically individuals of the most economically active ages who emigrate, forgone earnings and forgone own-production may outweigh the effect of fewer mouths and more remittances, potentially causing poverty to rise for those left behind, even if the migrants themselves become better off.

The biggest difficulty in estimating the impacts of emigration is posed by selectivity issues. A common research strategy in this literature is to use household survey data from the sending country to compare households where some members have emigrated to those where no one has emigrated. Such comparisons are complicated by a triple- or quadruple-selectivity problem: first, house-

We thank the Government of Samoa for permission to conduct the survey, the Samoa Bureau of Statistics for implementing the survey, and most especially the survey respondents. We also thank John Strauss and three anonymous referees for suggestions that greatly improved the paper. Financial support from the New Zealand Department of Labor and Marsden Fund grant UOW0503 is gratefully acknowledged. The views expressed here are ours alone and do not necessarily reflect the opinions of the World Bank, the New Zealand Department of Labor, or the Government of Samoa. This paper was initially written while Stillman was visiting the Department of Economics at the University of Zurich. Their wonderful hospitality was much appreciated. Contact the authors at dmckenzie@worldbank.org.

holds self-select into emigration; second, in some emigrant households everyone moves (and thus such households are almost never included in surveys in the sending country) while other emigrant households leave some members behind; third, some emigrants choose to return home, so their household may (wrongly) be considered as not affected by emigration. And fourth, if researchers wish to examine how the impact varies with duration since migration, they also face selectivity into not just whether, but when households emigrate.

In this paper, we set out how these selectivity issues arise and their implications for existing migration studies. We then address these selectivity problems by using new survey data designed specifically to take advantage of a randomized lottery that determines which applicants to the oversubscribed Samoan Quota (SQ) may immigrate to New Zealand. These data allow us to compare incomes and poverty among left-behind members in households in Samoa that sent SQ emigrants with incomes and poverty rates of similar households that were unsuccessful in the lottery. This random lottery solves the problem of self-selection into migration. The SQ policy rules control who can accompany the principal migrant, enabling us to address the second selectivity problem. Finally, our survey includes a module on return migration, allowing us to address the third selectivity problem. Differences among migrants in when they win the ballot lottery allow us to also explore duration effects and address this fourth selectivity issue.

We find that emigration reduces poverty among remaining members in the migrant-sending households. Although our sample is quite small for examining duration effects, we also find suggestive evidence of duration-dependent heterogeneity, with the point estimates suggesting that consumption and income fall relative to the first year effects as more time is spent abroad. This occurs because remittances and agricultural income decline with the duration since emigration and increases in household labor earnings with duration are not enough to offset this. We also demonstrate that failing to control for both selection into migration and selection into whether the whole household migrates leads to biased estimates and incorrect inferences on the impact of emigration on households left behind.

This is the first paper to use experimental data from the Samoan Quota lottery. In related work, we have looked at short-run (1 year) effects of a similar (but newer and smaller) migration program in Tonga (McKenzie, Gibson, and Stillman 2007; Gibson, McKenzie, and Stillman 2011). This paper builds significantly on our earlier work both methodologically and substantively. From a methodological viewpoint, this paper clearly lays out the additional selectivity issues that the existing literature has not fully addressed and provides guidance for both experimental and nonexperimental attempts to look

at the impacts of migration. Since the literature to date has not recognized all these dimensions of the selectivity problem, we believe the clear methodological framework provided here will be of interest and value to a range of other migration studies.

From a substantive viewpoint, the paper provides the first medium-term experimental estimates of the impact of migration—the impacts here are measured within 6 years of the eligible household members moving to New Zealand, and the first estimates that allow for duration-dependent heterogeneity while addressing selectivity. There are a number of theoretical reasons why the impact of migration on sending households is likely to vary with the duration of migration, and there are indeed reasons to believe that not just the magnitude, but also the sign, of any effects may differ in the short and medium term. Our results for Tonga are not able to examine this issue since data from there cover only one cohort of migrants. Our findings here show that allowing for this type of heterogeneity may be important in practice.

International migration is most prevalent and has the largest effects in small island countries, so this paper provides new evidence of the impact of migration in a country setting where it is of first-order importance. One of the most often discussed critiques of randomized experiments is that they are typically not replicated in other countries, raising concerns of external validity. This paper effectively doubles the set of countries for which experimental evidence on the impacts of migration is available. The fact that we find migration reducing poverty in Samoa contrasts with the findings in McKenzie et al. (2007) that migration increased poverty in the short run in Tonga, which we attribute to different intrahousehold selectivity into migration in the two countries, and potentially to more experience dealing with and preparing for migrant absence in the Samoan case.

II. Challenges to Understanding the Impact of Migration: Triple or Quadruple Selectivity and Duration-Dependent Heterogeneity of Impacts

There are now a sizable number of studies that aim to answer the question, "What is the impact of engaging in international migration on household incomes and poverty in sending countries?" If emigration purely resulted in an exogenous increase in income for the remaining members via remittances, the sign of the expected impact would be trivial. However, emigration can have a large number of other impacts on sending households. Most obviously, an absent migrant earns no domestic wage and provides no time inputs into market and household production. While economic theory suggests via revealed preference that the net impact on a well-informed household is positive for it to decide to migrate, it provides little guidance as to the magnitude of this effect

or how this gain will be distributed between the migrant and the remaining family members.

Indeed our previous work on migration from Tonga has found very large gains to the migrants themselves (McKenzie, Gibson, and Stillman 2010) but evidence that the family members left behind are, in the short run at least, worse off when some members migrate (Gibson et al. 2011). The net effect of migration on the remaining members of sending households is therefore a priori unclear and hence an empirical issue. However, there are several challenges to estimating this impact that the existing literature has largely ignored. We outline these challenges here and discuss why the existing literature is unable to overcome them. Then, in the next section, we discuss our approach for producing unbiased estimates of the impact of emigration.

A. *Triple Selectivity*

Assume for the moment that sending members abroad has a homogeneous impact β on the per capita household income of the remaining household members. The goal of many papers in the literature is to estimate this causal effect.[1] The standard approach is to begin by specifying a linear regression model for household i, relating per capita household income (or any other related outcome of interest), Y_i, to whether or not that household engages in international migration, M_i, and a set of observed (exogenous) characteristics of the household, X_i:

$$Y_i = \beta M_i + \boldsymbol{\gamma}' X_i + \varepsilon_i. \tag{1}$$

The standard concern is that self-selection into migration leads to biased estimates. In particular, one worries that there are unobserved attributes, such as personality type, entrepreneurial ability, and ambition, that are correlated with both the decision of the household to send migrants and the income that the household earns. That is, we are concerned that

$$E(M_i \varepsilon_i) \neq 0. \tag{2}$$

The existing literature has focused on trying to address this first form of selectivity using a variety of nonexperimental methods. This includes assuming selection on observables (Adams 1998), estimating parametric selection correction models (Barham and Boucher 1998; Acosta, Fajnzylber, and Lopez 2007), and using propensity-score matching (Esquivel and Huerta-Pineda 2007) and in-

[1] Some studies in the literature instead use receipt of remittances. This raises additional selectivity issues in terms of which migrants send remittances and how much they send (see McKenzie 2005), and so we prefer to focus on the broader overall impact of migration.

strumental variables (IV) methods (McKenzie and Rapoport 2007; Brown and Leeves 2011). However, it is easy to question the identification assumptions underlying these nonexperimental approaches. A number of recent papers show that migrants self-select in terms of both observables and unobservables (Akee 2010; McKenzie et al. 2010), in which case methods, like ordinary least squares (OLS) and matching, that assume selection on only observables produce biased results. Similarly, selection correction methods in the Heckman tradition rely on stringent functional form and distributional assumptions and dubious excludability restrictions. For example, Acosta et al. (2007) and Barham and Boucher (1998) assume that household assets predict selection into migration but do not directly affect earnings or labor force participation. Yet, these assets could be used to help finance business activities or themselves could be the result of labor earnings, so they are unlikely to be a valid instrument.

Similarly, many papers using IV methods rely on current migration networks as an instrument, which is subject to concerns about whether there are other excluded variables at the community level that also affect migration and outcomes of interest. For example, a recent community weather shock, such as a drought, may lead to both increased migration and a reduction in agricultural income in the community, so an empirical correlation between emigration and poverty would be a misleading estimate of emigration's impact. Historic networks are less subject to concerns about recent shocks but still need to rely on a plausible story of why networks exogenously formed in one location and not in another, such as the pattern of development of the railroad system in Mexico, as used by Woodruff and Zenteno (2007) and McKenzie and Rapoport (2007).

Moreover, the selection issue raised by equation (2) is only one of the three sources of selectivity that make it difficult to estimate β. The second source is selectivity among the households that engage in international migration as to whether or not the whole household moves. To see how this matters, note that we only observe the per capita income in equation (1) for households in which some members remain after other household members emigrate. Define a selection indicator s_i for each household i so that $s_i = 1$ if we observe Y_i and $s_i = 0$ if we do not due to the whole household moving. Then, rather than being able to estimate equation (1), all we can estimate is the following equation:

$$s_i Y_i = \beta s_i M_i + \gamma' s_i X_i + s_i \varepsilon_i. \tag{3}$$

Consistency of OLS estimation then requires assuming that

$$E(s_i M_i \varepsilon_i) = 0 \tag{4}$$

and

$$E(s_i \varepsilon_i) = 0. \tag{5}$$

Equation (4) is again the first form of self-selection considered, this time restricted to the group of households that are observed in the home country. The additional insight here is the need for equation (5) to hold. This requires that unobserved determinants of income are uncorrelated with whether a whole household leaves or not. It is important to note that this might not be true even if equation (4) holds. In particular, even if there is no self-selection of households into migration (which is unlikely), if households self-select in terms of whether or not the whole household moves, this will still generate biased estimates.

This formulation shows that if efforts are made to deal with the first selectivity, as is common in the existing studies, the results produced are still biased if there is self-selection among which households migrate en masse and which separate. The standard solution to account for this second form of selectivity is to explicitly model the process of selection into the sample and correct for this—either parametrically through the standard Heckman procedure or semi-parametrically. But this requires researchers not just to come up with an instrumental variable or convincing exclusion restriction that can be used to explain why some households migrate and others do not but also to find a second instrumental variable or exclusion restriction to explain which households migrate in full and which leave members behind.

This is not just a tall order, but is, in fact, impossible in most existing studies that only use data from the sending country (e.g., Adams 1998; Barham and Boucher 1998; Esquivel and Huerta-Pineda 2007). These studies, by definition, miss all households in which all members migrate (i.e., none are left in the sample population of the survey). Without data on the characteristics of these households, it is impossible to examine how they compare to those of households with remaining members or to model this selection process. Either policy rules that constrain migrant eligibility or data from the destination country on whole households that move are needed to model this selectivity.[2] Nev-

[2] An alternative approach is provided by Yang (2008), who provides some of the most convincing evidence on the impact of migration by using variation in exchange rate shocks experienced by migrant-sending households in the Philippines who send migrants to different destinations. Implicitly, he deals with this second selection issue by restricting the question of interest to the impact of migration on households that do not migrate en masse and then does not have to use nonmigrant households as controls, since the comparison is between different sets of partial migrant households. The main threat to his identification strategy is then one of selectivity into destination, which he addresses through a number of controls and sensitivity checks. This strategy is less amenable to cases where migrants move only to a single country, although if there is sufficient variation in economic conditions among destination cities, this could potentially be used as an instrument.

ertheless, in many cases it seems likely that the whole households that move are likely to be smaller and richer than households in which some households stay behind. Failure to eliminate such households from the comparison set of nonmigrant households will therefore bias toward finding that migration increases household size, reduces household per capita income, and increases poverty rates.

Finally, the third form of selectivity that raises a challenge for estimating the impact of migration on sending households arises from return migration. The magnitude of this effect will depend on type of migration and country context. While data on return migration rates are very limited, the evidence that does exist shows high rates of return migration in many contexts. For example, Gibson and McKenzie (2011) find among their study of top academic achievers that 38% of New Zealanders and 39% of Tongans who had ever migrated had also returned to their origin country. Dustmann and Weiss (2007) find that 45% of immigrants into the United Kingdom from the European Union, America, and the Middle East had remigrated within 5 years, although return migration rates were much lower for migrants from Africa and the Indian subcontinent. Gundel and Peters (2008) find that 15% of immigrants from outside the EU and 40% from within had remigrated within 20 years of migrating. Return migration rates will be higher still to countries where seasonal or temporary migration is the norm. There is evidence to suggest that return migrants are also self-selected, with Borjas and Bratsberg (1996) showing that return migration accentuates the type of initial self-selection, so that if migrants are positively selected from the origin population, return migrants are negatively selected from among the migrant group.

How return migrants should be treated in an analysis of the impact of emigration depends on what the parameter of interest is to the researcher. If the object of interest is to estimate the impact of a household currently having a migrant abroad, then return migrants should be classified as having $M_i = 0$ in estimating equation (3). However, this has consequences for ensuring that condition (4) is satisfied. In particular, it means that researchers must come up with an exogenous reason why some households have a migrant abroad who has not returned while others do not. Simply modeling the decision to engage in migration or using instruments, such as migration networks, that predict the initial migration decision will then no longer be enough—either return migration will need to be separately modeled or instruments that explain both the decision to migrate and the decision to stay abroad will be needed.

An alternative approach is to treat the parameter of interest as the impact of ever having engaged in international migration. Then, households with return migrants should be treated as having $M_i = 1$ in estimating equation (3). Stud-

ies that focus on the direct impacts of remittances typically do not treat return migrant households in this way (e.g., Adams 1998; Acosta et al. 2007; Adams, Cuecuecha, and Page 2008). However, research that recognizes the whole host of channels through which migration affects sending households (e.g., McKenzie and Rapoport 2007) often acknowledges that migration can continue to have impacts on households, for example, through repatriated savings and through knowledge and skills gained abroad, and so treat return migrant households as participating in migration. In this case, this third form of selection does not present additional challenges for estimation from the first two forms—it just requires that households with return migrants be correctly classified as migrant households.

B. Duration-Dependent Heterogeneity of Impacts

The discussion above assumes, in common with most of the literature, that participating in international migration has a constant effect on all households. If this assumption is violated and the impact of migration varies across households, then what is actually estimated is an average effect of some form—either the average effect from OLS in the absence of selectivity, or the marginal average treatment effect (MTE) or local average treatment effect (LATE) when methods to deal with selectivity are used. There is considerable discussion in the treatment effects literature as to when and whether these MTE and LATE estimates are parameters of interest for research or policy. We discuss these issues in the context of our estimation later in this paper.

However, we wish to point out a more fundamental issue when it comes to estimating the impact of migration on incomes and poverty in sending households. This is that the impact of migration on sending households is likely to vary with the duration of migration, and there are indeed reasons to believe that not just the magnitude, but also the sign, of any effects may differ in the short and medium term. For example, the short-term impact of migration may be negative as households lose the domestic income that the migrating members normally generated and perhaps have fewer assets to work with due to the costs of financing migration. It may take migrants some time to start paying off their moving costs and to earn enough to start sending remittances. However, in the medium term, this impact may be positive as the left-behind household members adapt to their new circumstances and receive greater remittances from migrants.

On the other hand, there is a debate in the literature as to whether remittances decay—that is, as to whether the amount of remittances received falls with duration abroad. If this is the case, then the short-run impact may be more positive than the longer-run impact. In either case, estimating the average effect

of migration over all households therefore gives an effect that might be accurate for at most some point between the short and medium terms, and it could miss most of the impacts of migration.

As a consequence, researchers should ideally move from the simple specification in equations (1) and (3) toward allowing the impact of migration to vary with the duration abroad. For example, if we let t_i be the number of years since emigration, we might be interested in estimating:

$$s_i Y_i = \beta s_i M_i + \lambda s_i M_i \times t_i + \gamma' s_i X_i + s_i \varepsilon_i. \quad (6)$$

Then the impact of having a household member abroad for t_i years is $\beta + \lambda t_i$. However, consistently estimating equation (6) requires an additional assumption on top of the selectivity assumptions raised previously. This requires that there can be no selectivity in terms of how long members have been abroad, that is, that

$$E(s_i M_i t_i \varepsilon_i) = 0. \quad (7)$$

Return migration is one reason why such selectivity could arise. However, even in the absence of selectivity into return migration, this assumption will be violated if the characteristics of households that sent migrants say 2 years ago differ from the characteristics of those households that sent migrants 5 years ago, that is, if there is selectivity not just in whether a household engages in migration but in when it does. Business cycle effects are one reason this assumption could be violated—the types of households that send migrants during a recession may differ from the types of households that send migrants during a boom. Researchers attempting to estimate (6) therefore face a quadruple-selectivity problem—selection into migration, selection among migrants as to whether the whole household migrates, selection into return migration, and selection into current duration abroad. Attempting to model these four forms of selectivity and control for them in a nonexperimental way poses an extreme challenge that most research designs are unlikely to be able to meet. We discuss next how the migration lottery we study allows us to address each of these factors.

III. Using an Emigration Lottery among Samoans to Address These Selection Issues

A. Background on Samoan Emigration

The country of Samoa consists of four inhabited islands in the South Pacific. The population of 180,000 is predominantly rural, with two-thirds of the labor force employed in agriculture. GDP per capita at market exchange rates is ap-

proximately US$2,000 (Vaai 2007), similar to Guatemala, Indonesia, and Morocco. In common with many small island nations, emigration and remittances are very important for Samoa. There are approximately 100,000 Samoa-born living overseas. Slightly over 50,000 of these emigrants live in New Zealand, with Australia, American Samoa, and the continental United States (including Hawaii) being the next most important destinations, each with approximately equal numbers.

Sizable migration from Samoa to New Zealand began during the 1960s and 1970s, with Samoans arriving on 3-month visas to take up work opportunities. After their permits expired, many stayed in New Zealand since the return provisions of the visas were not actively policed due to the excess labor demand at the time. But in the recession that followed the first oil crisis, labor demand fell sharply and, starting in 1974, "dawn raids" were launched to deport alleged overstayers. Since New Zealand had administered Samoa (then known as Western Samoa) under League of Nations and United Nations mandates from 1920 until 1962, the citizenship status of Samoans was uncertain, and a case was taken to the British Privy Council. In 1982, the council ruled that all Samoans born between 1924 and 1948 were British subjects and that when New Zealand citizenship was created in 1949, these Samoans and their descendants had also become New Zealand citizens. In response, the New Zealand government passed the Citizenship (Western Samoa) Act 1982 to overturn that ruling, restricting citizenship only to those already lawfully in New Zealand.

B. The Samoan Quota

However, as a compensation for this limitation on labor mobility, a "Samoan Quota" was agreed to as part of a Treaty of Friendship. This quota allows a specified number of Samoans to be granted New Zealand permanent residence annually, in addition to those entering New Zealand under normal immigration arrangements. The quota has been set at 1,100 places per year since 2002, and any Samoan citizen aged between 18 and 45 is eligible to register. The registration is free and many more applications are received than the quota allows, so a random ballot is used to deal with this oversubscription.[3] The number of registrations varies between 5,000 and 7,000 per year. Since the quota of 1,100 applies to the total of primary applicants and the secondary migrants who are eligible to immigrate with them, it represents about 400 family groups. Over the

[3] The random ballot was introduced in 1999. Prior to this, decisions were made on a first come, first served basis. The ballot was drawn manually up until 2003, from which point a computer was used to select randomly among registrations.

last 3 years, 1,201 out of 19,326 registrations were drawn in the ballot, representing odds of about 6%.

Permanent and long-term arrivals from Samoa to New Zealand average only 2,000 per year, while settlement migrants from Samoa average 500 per year into Australia and 200 per year into the United States over 2002–8, mostly through family reunification policies. It is thus clear that the Samoan Quota is a major channel for settlement emigration out of Samoa, accounting for approximately 40% of all emigration and the vast majority of emigration through a channel other than family reunification.

Once an applicant is selected in the random ballot, they must provide evidence of a job offer in New Zealand (unskilled jobs suffice) within 6 months in order to have their application to immigrate approved. These job offers are increasingly arranged by large employers visiting Samoa after the annual ballot results are announced. Once proof of a job offer is filed along with a residence application, it typically takes 3–9 months for an applicant to receive a decision, and they are then given up to 1 year to move. If they are successful, their immediate family (spouse and dependent children up to age 24) can also move to New Zealand with them. This rule specifying which family members can and cannot accompany the successful migrant, coupled with the random selection among Samoan Quota applicants, is key to being able to address the selectivity issues raised in the previous section.

Conceptually, we can estimate the impact of migration on family members left behind by comparing outcomes for the group of households in Samoa that sent SQ emigrants to those for the group with unsuccessful ballots who would not be eligible to move their entire household to New Zealand had their principal applicant been chosen in the ballot. Some households who win the lottery do not end up migrating, so we use an IV approach whereby we instrument whether former members of the household have migrated via the SQ with whether any were ballot winners. We discuss this approach in more detail in Section III.D. We next discuss the data that have been collected to allow us to implement this estimation approach.

C. The Samoa Labor Mobility Survey

The data used in this paper are from the Samoa Labor Mobility Survey (SLMS), which was designed by the authors and implemented by the Samoa Bureau of Statistics in late 2008. The survey is based on a self-weighting sample of 4,395 individuals from 622 households in 90 villages, drawn from all regions of Samoa. Out of these surveyed households, 78 had current or former members who were successful applicants to the Samoan Quota in the previous

5 years.[4] A further 121 households contained individuals that had applied at least once to the Samoan Quota in the previous 5 years and had never been successful. The remaining 423 households in the sample did not report having either former members who were now SQ emigrants or having current members who had applied to the SQ in the previous 5 years. These households that did not apply to the SQ are not good controls for the migrant households and are therefore not used in the experimental estimates in this paper. We do use this group to demonstrate the bias caused by ignoring selection in migration when estimating the impact of emigration on household members left behind.

The SLMS measures both income and expenditure of the responding households. Five income components are considered: labor earnings; net (i.e., also taking account of outbound) remittances of both money and goods over the previous 6 months; net returns from sales of fish, crops, livestock, and handicrafts; the value of own-produced or own-captured food consumed by the household; and other income from investments, pensions, rentals, and so forth. Household expenditures are recalled over the previous week, month, or 6 months, depending on the particular item, and an estimate of household consumption is formed from the sum of cash expenditures and the value of own-produced or own-captured food consumed by the household. All of the income and expenditure components are adjusted to an annual basis.

We use poverty standards that are based on existing poverty lines set for Samoa from the 2002 Household Income and Expenditure Survey. The higher "basic needs" poverty line had a value of ST$37.49 per adult equivalent per week, and 20.3% of households in Samoa had consumption expenditures below this level in 2002. The food poverty line, which was the required expenditure just for a minimum diet providing 2,200 kilocalories per adult per day, was calculated as ST$24.68 per week in 2002 (with 7.6% of households below this line). We used the Samoa consumer price index to update these poverty lines to October 2008 annual values of ST$2,962 and ST$1,850, which is equivalent to US$1,007 and US$663 per adult equivalent per year.[5]

In addition to these two objective poverty standards, the SLMS also asked the principal applicant in each ballot loser household and the adult aged 18–45 with the next birthday in ballot winner households about subjective poverty, using a 10-rung Cantril ladder question: "Please imagine a 10-step lad-

[4] We decided to focus on migration from the 2003–8 ballots to limit issues in regard to the accuracy of recalling past experiences and because we also have access to limited administrative data on SQ migrants from these ballots.

[5] The average exchange rate during the period of the survey was 2.94 Samoan Tala per US dollar. In the absence of any nutritional-based adult-equivalence scales we assume that children count as 0.5 of an adult.

der where on the bottom, the first step, stand the poorest people and on the highest step, the tenth, stand the richest. On which step are you today?"

D. Estimating the Impact of Migration through the Samoan Quota

The Samoan Quota enables us to address the triple- and quadruple-selectivity issues more credibly than existing studies. Consider first the problem of estimating the overall average impact of migration on household income per capita in Samoa, as in equation (3). The random selection among SQ applicants provides a means of overcoming the first selectivity issue—that households self-select into migration. Households self-select into whether or not a member applies for the SQ, but among these applicants whether or not they can migrate is random. The SQ policy rules provide a means for overcoming the second source of selectivity. The rules specify which individuals can migrate with a principal applicant—the spouse and dependent children. Thus, we restrict the sample further to drop households in which all members would be eligible to move to New Zealand if they had a successful ballot. This involves dropping approximately 22% of the 121 unsuccessful ballot households in our sample, who would move their whole household to New Zealand if they had won the SQ lottery. Then, equation (5) is satisfied for this subsample of ballot applicant households where someone would remain even if a household member was successful in the SQ lottery.

An additional potential issue could be that remaining household members from ballot-winning households move into new households when some members leave. This raises the concern that there are individuals in the migrant households who would not be there if not for migration. We examine this in Section IV.A and do not find any significant impact of migrating on household size or of migration on the composition of individuals ineligible for migration. But small sample sizes prevent us from ruling out this concern completely. The concern would then be that household resources change in the migrant-sending households due to the contributions of new household members rather than due to the direct effects of migration on those left behind. Such a change would still be an effect of migration, but just through a different channel. The available data and qualitative work from Samoa suggest that this channel is not a large one.

If all households that won the SQ lottery sent a migrant to New Zealand, then equation (4) would also be satisfied in this subsample. However, not all lottery winners moved to New Zealand—some may have changed their minds, others may still have been in the process of moving, while others may have been unable to find a job or failed another of the immigration requirements. In our sample of 78 households with ballot winners, 29 households (37%) did not

have a member who had yet moved to New Zealand via the SQ, although 11 of these from the most recent SQ ballots were in the process of moving to New Zealand. To address this drop-out bias, we employ the standard strategy in the experimental literature of using assignment to the treatment as an instrumental variable for the treatment itself. In other words, our instrumental variable is a dummy for whether a current or former household member had a successful ballot, whereas the treatment variable is whether someone from the household ever moved to New Zealand via the SQ. Randomization ensures that success in the ballot is uncorrelated with unobserved individual attributes that might also affect outcomes among the stayer household members and success in the ballot also strongly predicts migration.

What about the third source of selectivity, arising from return migration? In our sample, none of the 78 ballot winner households contained a member who had migrated to New Zealand after being selected in the five previous ballots but subsequently returned to live in Samoa. Our special module that was designed to capture return migration did find two former SQ migrants in sample households, but both had originally migrated in 2001 and hence were not in our analysis window. Hence, we were not required to deal with the impact of return migration in our analysis.

Restricting the sample to SQ ballot entrant households who would not all move if a member won the lottery and using success in the lottery as an instrument for migration therefore enables us to consistently estimate the impact of migration. If the impact varies across households, the impact we identify is a local average treatment effect (LATE), which in our case is also the average treatment effect on the treated, since none of the households losing the lottery can migrate through the SQ and likely do not have access to other migration channels (Angrist 2004). Some of the recent debate about the growing use of randomized experiments in development economics centers on whether the LATE is a parameter of interest in many policy experiments (e.g., Deaton 2010; Imbens 2010). In our case of a migration lottery, we believe that this parameter is clearly of policy interest, as it is an unbiased estimate of the impact of migrating for a Samoan household that applies to migrate through the migration policy being offered and does migrate if it is chosen. Thus, this is precisely the information needed to evaluate the development impact of this migration policy. Moreover, since the Samoan Quota is the main nonfamily migration channel out of Samoa, it is also the effect of interest for a large share of Samoan households with migrants.

The impact of migration is then measured for the remaining household members of SQ winners. These individuals are typically working-age and older adults who are either the parents or the siblings of the principal applicant, along

with children who are often their nephews and nieces. Specifically, 40% of household members left behind are under 18 and are mostly nephews and nieces of the principal applicant and spouse, 39% are working-age adults and are mostly the siblings of the principal applicant and spouse, and the remaining 22% are older adults who are mostly the parents, aunts, and uncles of the principal applicant and spouse. Many migration policies worldwide allow migrants to bring their spouse and children, while making it difficult for them to bring other family members, so these remaining members are likely to be similar to the remaining family members seen in many other countries where permanent emigration is common.

The final sample size for our experimental estimates is thus a sample of 166 households, consisting of households with an SQ ballot winner and those with an SQ ballot loser who would have members remaining even if the ballot had been successful. Although one would always like larger samples, there are two important points to note in regard to this. First, the sample size we have for migrants is not that different from what is often obtained in studies using large nationally representative samples from other countries.[6] Second, and more important, the vast majority of households in large nationally representative household surveys are not similar enough to migrant households to form a valid counterfactual.[7] Thus, papers that seek to estimate the impact of migration, whether through experimental or nonexperimental methods, should be eliminating most nonmigrant households from the analysis. In our case, the lottery provides us a way of ensuring that we have an appropriate counterfactual for our migrant households and yields a (small) sample size that is no different than what one would typically obtain using nonexperimental methods appropriately.

The initial impact of migration that we estimate is an average over households whose members have been abroad for varying amounts of time. At the time of our survey, the sampled Samoan households with SQ emigrants in

[6] For example, the Mexican Family Life Survey interviewed households in 2002 and 2005. Rubalcava et al. (2008) use this to look at selection into migration by health status. Despite having a large urban sample of 3,817 young adults in a high migration country, there are only 144 urban migrant individuals over the 3-year period. Adams et al. (2008) have a nationally representative sample of 4,000 households from Ghana, of which only 210 receive international remittances (and certainly an even smaller number would have sent migrants in the last 2 or 5 years). Other specialized surveys in the literature are often on a similar scale. For example, Barham and Boucher (1998) have a total sample of 152 households (including both migrant and nonmigrant households) in their study of migration in Nicaragua.

[7] For example, in McKenzie et al. (2010) we use propensity-score matching with the Tongan Labor Force Survey and, even with a relatively small set of characteristics to match on, find that only 354 of the 4,043 observations are in the common support.

New Zealand had a mean (median) time abroad for their former household members of 3.4 years (3 years). Of the households in our sample, 37% were interviewed 1–2 years after eligible household members had immigrated to New Zealand, 31% were interviewed 3–4 years postmigration, and 33% were interviewed 5–6 years postmigration. Thus, our sample covers both the short- and medium-run impacts of migration.

As discussed above, there are theoretical reasons to suspect that the effect of migration may vary with duration; thus we also estimate equation (6). We do this by instrumenting the interaction between duration abroad and migrant status with the interaction between duration abroad and ballot success. Identification of the duration effect requires overcoming the fourth selectivity discussed in Section II, that there is not selection among the timing of migration. The Samoan Quota provides us with a plausible reason why households sent migrants at different points in time—they just happened to have their ballot drawn in different rounds. We would still be concerned about selectivity if the characteristics of households applying (and thus of those winning) in different years differ significantly. Our sample sizes are too small to examine year-to-year variation in household characteristics, but, as a simple check, we compare the characteristics of ballot winner households with members abroad for less than the median duration to those abroad for more than the median duration. The results suggest that there is little selection into when individuals entered the SQ ballot over the years considered here.

E. Verifying Randomization

We first test whether the lottery correctly randomizes households into a treatment and a control group by examining whether the households containing ballot losers are statistically different from the households containing ballot winners (both the emigrant-sending households and the noncompliers). As discussed above, attention is restricted to households in which some members would have stayed according to the age and relationship rules on which secondary applicants may accompany the principal applicant. Table 1 compares the ex ante premigration characteristics available from the survey for ballot winners to those who were unsuccessful, restricting the sample to individuals who had applied to the SQ in the previous 2 years. We restrict this test to this subset of our sample since the survey did not collect information on whether each household member at the time of the survey resided in the household when the SQ winners emigrated in the past. Differences between ballot winners and losers from further in the past may just reflect changes in household composition that occurred after the SQ migrants emigrated and were potentially caused by this subsequent change in household composition.

TABLE 1
TEST OF RANDOMIZATION

Stayer Household Characteristics (n = 90)	Successful Ballot	Unsuccessful Ballot	t-test p-value
Size of the stayer household	5.8	5.2	.367
Number of adults 18–45 among stayers	2.3	2.8	.159
Number of children <18 among stayers	2.6	2.2	.446
Number of adults >45 among stayers	1.3	1.4	.435
Proportion of adults 18–45 who are female	.46	.39	.267
Mean age of stayer adults	31.8	27.6	.000
Mean years of education of stayer adults	13.2	13.9	.229
Located in Apia	.20	.25	.602
Located in Savai'i	.20	.17	.701

Note. Subsample of households from only 2007 and 2008 ballots. Characteristics are measured 1–2 years after randomization.

Of the nine characteristics we examine in this table, the only significant difference at conventional significance levels is found for the mean age of stayer adults, with stayer adults in successful households older than those in unsuccessful households. Given the number of variables we are testing, this is consistent with a joint test indicating that the characteristics of successful and unsuccessful households are globally insignificantly different from each other.

F. Accounting for Multiple Entries into the Migration Lottery

An additional further complication in our context is that households could decide each year whether or not to enter the Samoan Quota ballot over the 2004–8 period. A possible concern then is that households which entered the ballot in more years would be more likely to end up with a winning ballot and may also differ in other dimensions from households which enter the ballot fewer times during the 5-year period we consider. In theory, this problem is easy to fix provided one has data on which lotteries a household entered—one can then simply add dummy variables for each lottery entered as additional control variables in equation (1), and then whether or not a household wins the lottery would be random conditional on which lotteries they entered.[8]

However, although we have data on which lotteries the ballot losers and noncompliers entered, we do not have this information for the winning ballots who migrated. The reason for this is that we were doubtful that the families of migrants would be able to supply accurate retrospective information on which lotteries the migrants had entered and not won prior to their winning ballots. We therefore must instead examine how large an issue this is likely to be and show that our results are robust to placing additional weight on the ballot losers who entered multiple times.

[8] See Abdulkadiroglu et al. (2011) for an application in the context of schooling lotteries.

TABLE 2
ARE HOUSEHOLDS IN MULTIPLE LOTTERIES DIFFERENT FROM THOSE IN ONLY ONE?

Characteristics of Unsuccessful Households (n = 121)	Applied Once	Multiple Applications	t-test p-value
Household size	7.7	7.8	.883
Number of adults 18–45	3.2	3.2	.871
Number of children <18	3.3	3.3	.897
Number of adults >45	1.1	1.2	.601
Proportion of adults 18–45 who are female	.44	.50	.251
Mean age of adults	29.6	30.0	.761
Mean years of education of adults	12.4	13.1	.086
Total household income	21,934	31,968	.042
Located in Apia	.27	.20	.474
Located in Savai'i	.18	.08	.238

Note. Out of 121 households, 25 applied to multiple lotteries between 2004 and 2008.

We were able to obtain administrative data from the New Zealand Department of Labor on the results of the 2005, 2006, and 2007 lotteries. Among the ballot winners, 67% had applied only once over this 3-year period, 25% had applied twice, and 7% had applied in all three periods. This compares to 72% of the ballot losers having applied only once over this 3-year period, 21% having applied twice, and 6% having applied three times. So, while it is the case that the ballot winners are more likely to have entered multiple times, the difference is not that great, and the majority of applicants had applied only once. This suggests that in our context accounting for multiple entry is unlikely to make much difference unless the characteristics of those entering multiple times differ dramatically from those who enter only once.

Next we compared the characteristics of the unsuccessful applicant households in our sample who applied to only one ballot during the 5-year period to those of the remaining (21%) unsuccessful households with multiple ballots (see table 2).[9] No significant differences were found in the demographic composition of households (size, age, or gender) or their location. However, adults in households that applied to multiple lotteries are slightly more educated, and these households have significantly higher incomes. Given that for the most part the characteristics of households entering multiple ballots are either the same or fairly similar to those entering only one ballot, there is little reason to believe that failing to account for this affects our results qualitatively. But, as a final check on the robustness of our results to multiple entries, we reweight our ballot loser sample to give more weight to the households that had entered multiple lotteries (i.e., we reweight them from being 21% of the sample to being

[9] Note that 21% of our sample had entered multiple ballots, as compared to 27% of all households in the administrative data. This difference is not significant; thus we use the administrative proportions in reweighting.

33% as for winners in the administrative data). The results from this exercise are presented in appendix table A1 and show that accounting for different rates of multiple entry among ballot winners and losers has no material impact on any of our estimates.

IV. Experimental Estimates of the Impacts on Income and Poverty

In this section, we present experimental estimates of the impact on income and poverty of sending household members to New Zealand under the Samoan Quota. As discussed above, here the age and relationship rules governing which secondary applicants can move with the principal applicant are used to identify and drop control group households where all members would have moved to New Zealand if they had a successful ballot. Because the exogenous covariates we can measure for households are at the time of the interview and not at the time the SQ migrant left the household, we do not control for any covariates except household location (four regions) in these regressions.

A. The Impact on Household Size and Composition

We begin in table 3 by examining the impact of emigration of some household members on household size and composition, since one immediate effect is that there are "fewer mouths to feed." These results are unweighted and thus indicate the change in household size for the average household. Emigration leads to a significant reduction in household size. The mean household among unsuccessful applicant households that are not entirely composed of individuals who would have migrated had the ballot been won has 8.2 people, and emigration is estimated to reduce this by 1.2 people. This is a 15% reduction in household size. Emigration leads to households having, on average, 0.9 fewer prime-age adults, which is a 24% decrease. There is no change in the number of older adults (>45 years), which is reassuring since they are not eligible to move as secondary applicants. The impact of the number of children is also insignificantly different from zero but has a large standard error.

Table 3, panel B, interacts the indicator for whether a household has had members emigrate to New Zealand with a continuous variable measuring how long ago these members first emigrated. Thus, the main effect in this table shows the immediate impact of emigration on household size, while the interaction term shows how this impact depends on how long ago the migrant left the household. This panel also shows the total estimated impact on households where individuals emigrated 1, 3, and 5 years ago. Most of the coefficients in the specification are estimated quite imprecisely, making it difficult to say anything about how the impact of migration on household size varies with the duration of migration.

TABLE 3
IIMPACT OF MIGRATION ON HOUSEHOLD SIZE AND COMPOSITION

	Total Household Size	Adults Aged 18–45	Children Aged under 18	Adults Aged over 45
Mean for unsuccessful stayer households	8.23	3.53	3.31	1.39
A. Estimates pooling all households:				
Impact of migration	−1.24*	−.85**	−.21	−.18
	(.72)	(.33)	(.54)	(.19)
Sample size	166	166	166	166
B. Estimates allowing for duration effects:				
Impact of migration	−.51	−.60	.52	−.43
	(2.25)	(1.04)	(1.68)	(.61)
Added impact of each year in New Zealand	−.21	−.07	−.21	.07
	(.51)	(.23)	(.38)	(.14)
Impact of 1 year in New Zealand	−.72	−.67	.31	−.35
	(1.77)	(.81)	(1.32)	(.48)
Impact of 3 years in New Zealand	−1.14	−.82**	−.11	−.21
	(.89)	(.41)	(.66)	(.24)
Impact of 5 years in New Zealand	−1.56**	−.96***	−.53	−.07
	(.69)	(.32)	(.52)	(.19)
Sample size	166	166	166	166

Note. These results are from instrumental variables regressions where migration is instrumented with the SQ ballot outcome. The only control variables are indicators for the location of household in Samoa. Standard errors are in parentheses.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

B. The Impact on Total Household Resources

Table 4 examines the impact of emigration on household total resources and the composition of household income. Since emigration changes household size and demographic composition these measures are not necessarily good proxies for individual welfare. However, they do show the ways in which households adjust their economic activities in response to emigration. Again, these results are unweighted and thus indicate the change in total resources for the average household.

The point estimates suggest that households which sent emigrants now have larger total household income and consumption than households who were unsuccessful in the lottery, but large standard errors on these estimates make the estimates statistically insignificant. The results do show a change in the composition of household income. Income from agricultural production and

TABLE 4
IMPACT OF MIGRATION ON TOTAL RESOURCES AND COMPONENTS OF HOUSEHOLD INCOME

	Total Income	Household Labor Earnings	Agricultural Income	Subsistence Income	Net Remittances	Total Consumption
Mean for unsuccessful stayer households	22,860	14,377	443	3,785	2,227	25,143
A. Estimates pooling all households:						
Impact of migration	1,961	−1,566	1,087***	528	1,679**	2,700
	(3,585)	(3,194)	(373)	(708)	(738)	(2,870)
Sample size	163	163	163	161	162	162
B. Estimates allowing for duration effects:						
Impact of migration	4,681	−7,639	3,530***	2,282	4,575**	6,955
	(11,329)	(10,141)	(1,280)	(2,255)	(2,303)	(8,849)
Added impact of each year in New Zealand	−773	1,723	−694**	−500	−831	−1,218
	(2,523)	(2,259)	(286)	(503)	(514)	(1,982)
Impact of 1 year in New Zealand	3,908	−5,916	2,836***	1,782	3,744**	5,737
	(8,918)	(7,981)	(1,007)	(1,774)	(1,813)	(6,962)
Impact of 3 years in New Zealand	2,362	−2,470	1,448***	783	2,083**	3,300
	(4,479)	(4,003)	(504)	(889)	(914)	(3,513)
Impact of 5 years in New Zealand	816	976	61	−216	422	863
	(3,394)	(3,030)	(390)	(675)	(706)	(2,763)
Sample size	163	163	163	161	162	162

Note. These results are from instrumental variables regressions where migration is instrumented with the SQ ballot outcome. The only control variables are indicators for the location of household in Samoa. Standard errors are in parentheses.
** Significant at the 5% level.
*** Significant at the 1% level.

income from remittances are significantly higher, by 245% and 75%, respectively, relative to unsuccessful households, while household labor earnings are lower (but not significantly so).

Table 4, panel B, then examines duration effects. We find that the impact of migration on agricultural income and remittances declines with duration, although the interaction term is insignificant for remittances. One explanation for this might be that emigrants work extra hard on the gardens of their sending families before they leave, in order to plant extra crops that can provide their families with income and food in the first year while the migrant is away. We have received reports of this occurring for workers participating in a new seasonal worker program. The effect of this would then dissipate after the first

year's crops are harvested.[10] An alternative explanation might be that the increase in remittances received in the first year is being used to purchase fertilizer and other agricultural inputs to increase production, and then as remittances appear to fall with duration, less of this occurs. Our survey does not allow us to test these hypotheses.

Conversely, the results suggest that the initial negative impact of migration on household labor earnings may decline as migrants spend more time abroad, possibly as remaining household members adjust to the absence of the migrants and/or as they adjust their labor supply in response to the declining remittances and agricultural income. The point estimates suggest that this response does not offset the decline in the other income components relative to the initial impact as emigrants spend more time abroad. But our small sample size among the treatment group means that we cannot reject that income and its components other than agricultural income are level over time, despite the sizable economic significance of the point estimates.

C. *The Impact on Per Person Resources*

Table 5 examines the impact of emigration on per person resources. We now weight the estimates by household size and hence the results indicate the change in per person resources for the average individual. These results show that left-behind household members are better off in comparison to members of households with lottery losers. Average consumption is approximately 17% higher in per adult-equivalent terms and income is approximately 23% higher (although neither is statistically significant).[11] Since the changes in income and consumption are similar, this suggests that these changes associated with emigration and remittances are being viewed as shocks to permanent income by the left-behind household members. There is some weak evidence that these gains become smaller over time, both because household size is rising and because the income gains are declining. However, these results are also consistent with the impacts being independent of how long the emigrants have been gone.

[10] The dominant crop for both market sales and subsistence in Samoa is taro (*Colocasia esculenta*) which can be harvested for 6–14 months after planting.

[11] Note that in common with other surveys, we do not collect data on the intrahousehold composition of consumption. One might worry that the rise in per person expenditure is the results of the individuals who migrate being the ones who had the lowest shares of household resources. This does not seem plausible here—the migrants are prime-age adults and their children, and resources are generally shared fairly equally within households in Samoa. Moreover, in terms of income, the principal applicants in Samoa do not have significantly different earnings from other 18–45-year-old adults in the same household (with the point estimates suggesting that if anything they earned more than nonapplicant adults in the same household).

TABLE 5
IMPACT OF MIGRATION ON PER ADULT EQUIVALENT RESOURCES

	Income per Adult Equivalent	Log Income per Adult Equivalent	Consumption per Adult Equivalent	Log Consumption per Adult Equivalent
Mean for unsuccessful stayer households	3,533	7.93	3,983	8.15
A. Estimates pooling all households:				
Impact of migration	819	.17	977	.23
	(663)	(.19)	(607)	(.14)
Sample size	164	164	164	164
B. Estimates allowing for duration effects:				
Impact of migration	877	.27	1,507	.25
	(2,211)	(.64)	(2,016)	(.45)
Added impact of each year in New Zealand	−17	−.03	−155	−.01
	(508)	(.15)	(463)	(.10)
Impact of 1 year in New Zealand	860	.24	1,352	.25
	(1,722)	(.49)	(1,571)	(.35)
Impact of 3 years in New Zealand	827	.19	1,042	.23
	(816)	(.23)	(746)	(.17)
Impact of 5 years in New Zealand	793	.13	733	.21
	(660)	(.19)	(603)	(.14)
Sample size	164	164	164	164

Note. These results are from instrumental variables regressions where migration is instrumented with the SQ ballot outcome. The only control variables are indicators for the location of household in Samoa. Estimates are variance weighted by household size. Standard errors are in parentheses.

D. The Impact on Poverty

Finally, table 6 examines the impact of emigration on poverty. Again, we weight the estimates by household size and hence the results indicate the change in poverty for the average individual. The basic needs poverty rate among individuals living in households that sent Samoan Quota emigrants is 23 percentage points lower than for households with lottery losers. Since the poverty rate among individuals in unsuccessful households is 37%, this represents a 62% reduction in head count poverty. However, there is no measured effect of emigration on the food poverty rate that captures deeper poverty (with only 12% of the lottery loser households below this line), nor is there any effect on the poverty gap ratio at either poverty line.[12] This indicates that the main impact

[12] It is worth noting that given our small sample sizes, the impacts on threshold outcomes, such as poverty rates, will be especially sensitive to where the thresholds are drawn.

TABLE 6
IMPACT OF MIGRATION ON POVERTY AMONG REMAINING HOUSEHOLD MEMBERS

	Poverty Head Count Basic Needs Line	Poverty Head Count Food Poverty Line	Poverty Gap Basic Needs Line	Poverty Gap Food Poverty Line	Subjective Poverty Ladder
Mean for unsuccessful stayer households	.37	.12	.09	.03	5.44
A. Estimates pooling all households:					
Impact of migration	−.23**	−.01	−.03	−.01	−.29
	(.11)	(.08)	(.04)	(.03)	(.51)
Sample size	166	166	166	166	161
B. Estimates allowing for duration effects:					
Impact of migration	−.32	.19	.01	.03	−1.58
	(.38)	(.27)	(.14)	(.08)	(1.74)
Added impact of each year in New Zealand	.03	−.06	−.01	−.01	.38
	(.09)	(.06)	(.03)	(.02)	(.40)
Impact of 1 year in New Zealand	−.29	.13	.00	.02	−1.20
	(.29)	(.21)	(.11)	(.06)	(1.35)
Impact of 3 years in New Zealand	−.24*	.01	−.03	−.01	−.45
	(.14)	(.10)	(.05)	(.03)	(.64)
Impact of 5 years in New Zealand	−.19	−.11	−.06	−.03	.31
	(.11)	(.08)	(.04)	(.03)	(.52)
Sample size	166	166	166	166	161

Note. These results are from instrumental variables regressions where migration is instrumented with the SQ ballot outcome. The only control variables are indicators for the location of household in Samoa. Estimates are variance weighted by household size. Standard errors are in parentheses.
* Significant at the 10% level.
** Significant at the 5% level.

of having family members migrate to New Zealand was to lift individuals who were just below the basic needs poverty line out of poverty, with insignificant effects on the much smaller proportion of individuals who were poorer than this. The impact on the subjective poverty reported by an individual adult respondent in each household is negative but also statistically insignificant. Again, there is some weak evidence that any possible poverty reduction declines over time, but the years since migration term is neither significant nor the same sign across different poverty measures.

V. The Importance of Different Sources of Selection Bias

In this section, we examine how our experimental estimates compare to estimates that do not control for self-selection by households into emigration and the selection among emigrant households into those whether everyone moves

versus households that leave some members behind. This allows us to judge how large an effect these two forms of selection bias have on estimates of the impact on income and poverty of sending household members to New Zealand under the Samoan Quota. It is worth noting that both the direction and size of these two forms of selection bias are particular to the context of the Samoan Quota and may not be indicative of what would be found when examining the impact of emigration in other contexts.

A. Households with Left-Behind Members versus All-Move Households

We begin by examining the bias that results from failing to exclude unsuccessful households in which all members would move if the household had been successful in the migration lottery. We do this by reestimating the experimental regression models but including these "all-move" households as part of the control group. This replicates the situation in which a researcher has a convincing empirical strategy for controlling for self-selection into emigration but is unable to either identify which nonemigrant households would emigrate en masse or to model this selection process.

Table 7 presents the results from this exercise along with the original experimental estimates previously presented in panel A of tables 3–6. Because of the structure of the SQ policy rules, whole households that move are smaller than stayer households. Hence, failing to drop all-move ballot losers causes us to understate the fall in household size from migration. In this particular setting, this is the main bias that occurs from failing to control for this source of selection bias. On the other hand, failing to remove all-move households has no impact on our estimates of the impact of emigration on household income. This indicates that while all-move households have fewer members than the other households with lottery losers, they have similar levels of income. Hence, when examining household income, wrongly including the all-move households in the control group does not have a qualitative impact on the results. Turning to the remaining estimates, there is suggestive evidence that including all-move households leads to a downward bias in the estimated impact of emigration on per capita consumption and poverty rates, but none of these differences are estimated with enough precision to rule out the estimates being unaffected by this form of selection bias.

B. Selection into Emigration

We next examine the bias that results from failing to control for any source of selection on unobservables, either into emigration or among emigrant households, of whether to migrate en masse. Comparing these results to those in the previous section gives an indication of the importance of selection (on un-

TABLE 7

THE IMPORTANCE OF DIFFERENT SOURCES OF SELECTION BIAS

	Experimental Estimates (1)	Experimental Including All-Move Households (2)	Experimental Sample: OLS with Controls (3)	Full Sample: OLS with Controls (4)
Total household size	−1.24*	−.75	−.12	.29
	(.72)	(.72)	(.50)	(.28)
Adults aged 18–45	−.85**	−.53	.11	.2
	(.33)	(.34)	(.22)	(.13)
Children aged under 18	−.21	−.27	−.55	−.13
	(.54)	(.53)	(.36)	(.19)
Adults aged over 45	−.18	.06	.33***	.25***
	(.19)	(.21)	(.13)	(.07)
Total income	1,961	3,395	−3,278	364
	(3,585)	(3,599)	(2,284)	(1,372)
Household labor earnings	−1,566	−762	−5,005**	−3,275***
	(3,194)	(3,149)	(1,963)	(887)
Agricultural income	1,088***	1,097***	57	159
	(373)	(356)	(246)	(468)
Subsistence income	528	955	286	756***
	(708)	(682)	(462)	(244)
Net remittances	1,679**	1,982***	1,681***	2,733***
	(738)	(714)	(477)	(382)
Total consumption	2,701	2,239	−822	446
	(2,870)	(2,884)	(1,916)	(948)
Income per adult equivalent	819	888	−622	−133
	(663)	(665)	(410)	(273)
Log income per adult equivalent	.17	.18	−.14	.08
	(.19)	(.19)	(.12)	(.07)
Consumption per adult equivalent	977	529	−159	−78
	(607)	(616)	(364)	(202)
Log consumption per adult equivalent	.23	.14	−.02	.06
	(.14)	(.14)	(.08)	(.05)
Poverty head count, basic needs line	−.23**	−.20*	.002	−.04
	(.11)	(.11)	(.07)	(.04)
Poverty head count, food poverty line	−.01	−.01	−.02	−.06**
	(.08)	(.08)	(.05)	(.03)
Poverty gap, basic needs line	−.03	−.03	−.003	−.04**
	(.04)	(.04)	(.03)	(.02)
Poverty gap, food poverty line	−.011	−.006	−.004	−.03***
	(.03)	(.02)	(.02)	(.01)
Subjective poverty ladder	−.29	−.43	.28	−.07
	(.51)	(.50)	(.32)	(.18)
Sample size	166	199	199	615

Note. The first column repeats the complete results from panel A in tables 3–6. The second column presents the results when these models are reestimated with all-move ballot loser households also included in the sample. The third and forth columns present results from estimating OLS models with controls for the number of children, prime-age adults, and older adults in the households (except when household size is the outcome being examined), the percentage of adults in the household that are female, the average age of adults in the household, the highest educational attainment of adults in the household, household wealth, and the location of the household in Samoa. Here emigrant households are defined as any that receive international remittances from former family members. In the fourth column, SQ nonapplicant households are added to the sample. Standard errors are in parentheses.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

observables) into emigration on its own. Here emigrant households are defined as those receiving international remittances from former family members. This is the approach used to identify emigrant households in much of the literature since household surveys rarely ask about the current location of previous household members (e.g., Adams 1998; Acosta et al. 2007; Adams et al. 2008).[13] Appendix table A2 compares characteristics of the ballot loser households to households that did not apply for the Samoan Quota ballot (col. 2), to the subset of nonapplicant households that have an emigrant abroad (col. 3) and to households who were successful in the ballot but failed to move (i.e., noncompliers, col. 4). We also compare the subset of unsuccessful ballots that have an emigrant abroad as measured by remittance receipt to ones that do not (cols. 5 and 6).

There is clear evidence of positive selection into applying for the lottery: households that apply are more likely to be located in Apia (the capital), have higher education among adults, greater wealth, and larger household sizes. However, among ballot winners, there appears to be negative selection into migration—as seen by the fact that the noncompliers have significantly higher wealth than unsuccessful households and (insignificantly) higher education. Among unsuccessful ballots, those receiving remittances from former members, who are hence misclassified as emigrant households in these regressions, appear to be negatively selected, although again the evidence is quite weak in terms of statistically significant differences.

We then estimate OLS regression models (i.e., eq. [1] above) of the relationship between being an emigrant household and outcomes controlling for the number of children, prime-age adults, and older adults in the households (except when the regression outcome is household size), the percentage of adults in the household that are female, the average age of adults in the household, the highest educational attainment of adults in the household, household wealth, and the location of the household in Samoa.

The results from these models are presented in the third and fourth columns of table 7. Column 3 retains the same sample as column 2 but uses OLS rather than the experimental IV. This shows the effect of failing to control for selective migration among ballot winners as well as the effect of classifying households based on remittance receipt as opposed to a tightly defined criterion. Column 4 then uses the full sample of households, which induces additional bias by including as comparators households in which no one has applied to migrate.

[13] Seventy-one percent of unsuccessful ballot households receive remittances from former members versus 76% of SQ migrant households; hence this change in classification potentially has important impacts on the results.

Compared to column 2, we can see that failing to control for both forms of selection into emigration has larger effects on our results than failing to control for selection into en masse emigration. Both selection into migration among the ballot winners and selection into applying for the Samoan Quota matter. First, examining the impact on household size, selection into migration further exacerbates the downward bias caused by failing to exclude all-move households—consistent with the applicant households having larger household size than the nonapplicant households in appendix table A2. Using the specifications in columns 3 and 4, we would (incorrectly) find that emigration has no impact on household size and leads to an increase in the number of older adults in the household.

Failing to control for selection into migration also has large effects of our estimates of the impact on household income. Using these specifications we would find that emigration leads to a large decline in household labor earnings, no increase in agriculture income, but a larger increase in remittances. Combining these various impacts, we would now find emigration to have no impact on total household income, as well as no impact on per capita income and consumption, and a much smaller impact on basic needs household poverty.

Again, there is a lack of precision in many of these estimates, but overall the results strongly suggest that failing to control for self-section into emigration leads to a downward bias in the estimated impact of emigration on household income and a corresponding upward bias on the estimated impact on household poverty. Based on the evidence in appendix table A2, this bias is likely caused by negative selection into migration among successful ballots and into remittance receipt among unsuccessful ballots. For unsuccessful ballots, this is consistent with extended family members targeting remittances toward less well-off households or those that have experienced negative shocks, such as crop failures. It is uncertain to the cause of positive selection among noncompliers, but given that 11 out of 29 households in this group are in the process of moving to New Zealand, one possible explanation is that wealthier households find the moving process more time consuming (e.g., they may have more livestock to sell). While these findings are particular to the context examined in this paper, the results strongly suggest that the traditional approaches used in the previous literature examining the impact of migration on sending households are likely to produce biased results.

VI. Conclusions

The main difficulty in estimating the impacts of emigration on household members left behind is posed by selectivity issues. A common research strategy in this literature is to use household survey data to compare outcomes for

households in which some members have emigrated to outcomes for those where all members are currently in the sending country. Such comparisons are complicated by a triple-selectivity problem: first, households self-select into emigration; second, in some emigrant households everyone moves (and thus such households are almost never included in survey data on the sending country); third, some emigrants choose to return home, so their household may (wrongly) be considered as not being affected by emigration.

In this paper, we show how these selection problems invalidate the existing approaches used in most of the previous literature. Going forward, we recommend that nonexperimental studies model or instrument both selection into migration and selection as to which members migrate, and that where possible, they treat households with return migrants as still being affected by migration. In order to model the second form of selectivity, future studies will need to pay more attention to the policy rules and household behaviors that lead some households to send only some of their members abroad and other households to migrate en masse. In the absence of detailed institutional knowledge of policy rules that determine who can move, it will be important to use data on whole households that move (from destination country surveys) to determine which nonmigrant households should be dropped when comparisons are being made to households in the sending country with a migrant abroad.

Here we have addressed these selectivity problems by using survey data designed specifically to take advantage of a randomized lottery that determines which applicants to the oversubscribed Samoan Quota may immigrate to New Zealand. We find that migration reduced basic needs poverty among former household members. This contrasts with the short-term impact of migration for Tongans moving to New Zealand through a similar lottery program—in McKenzie et al. (2007) we find an increase in poverty for those left behind. Comparing the results, the difference stems from a much larger fall in household labor earnings in Tonga than in Samoa, and a fall in agricultural production in Tonga compared to a rise in agricultural income in Samoa. The increase in remittances is similar in both cases.

In Tonga, we found that the movers earned twice the weekly income as the stayers within a potential migrant household (Gibson et al. 2011), which was why household earnings declined so drastically when these individuals migrated. In contrast, fortnightly earnings for stayer 18–45-year-old adults in Samoa average 179 Tala, versus 211 Tala for the movers ($p = .32$).[14] Thus,

[14] This is based on analysis of earnings of potential movers and potential stayers in the ballot loser households containing both Samoan Quota applicants and individuals who would stay in Samoa had these applicants won the ballot.

Samoan households rely relatively less on the labor earnings of the potential migrants before migration and so suffer less opportunity cost of their absence in terms of these forgone labor earnings. It is also possible that the much longer history of migration through the SQ has given households more time to learn how to adjust to the absence of members, potentially through getting them to set up a lot of agricultural production beforehand. These differences highlight the importance of context in understanding the impacts of migration and of replicating these experimental studies in more locations where possible.

Finally, in addition to forming experimental estimates of the average impact of migration on left-behind household members, we estimated models that allow for duration-dependent heterogeneity in these impacts. While our sample size is small for precise estimation of such effects, we do find suggestive evidence that allowing for this type of heterogeneity may be important in practice. Our point estimates suggest that income among sending households decays as SQ migrants spend longer in New Zealand, with agricultural income and remittances declining with duration. While these results are only suggestive, they do point to a need for other studies with larger samples to seriously consider and model these duration effects rather than assuming a homogeneous impact of migration with time spent abroad.

Appendix

TABLE A1

ROBUSTNESS TO REWEIGHTING FOR MULTIPLE LOTTERY ENTRY

	Unadjusted	Reweighted
Total household size	−1.24*	−1.15
	(.72)	(.73)
Adults aged 18–45	−.85**	−.82**
	(.33)	(.33)
Children aged under 18	−.21	−.15
	(.54)	(.53)
Adults aged over 45	−.18	−.18
	(.19)	(.20)
Total income	1,961	1,505
	(3,585)	(3,751)
Household labor earnings	−1,566	−1,178
	(3,194)	(3,223)
Agricultural income	1,087***	1,073**
	(373)	(419)
Subsistence income	528	329
	(708)	(724)
Net remittances	1,679**	1,508*
	(738)	(784)
Total consumption	2,700	2,537
	(2,870)	(2,931)
Income per adult equivalent	819	718
	(663)	(695)
Log income per adult equivalent	.17	.13
	(.19)	(.20)
Consumption per adult equivalent	977	846
	(607)	(661)
Log consumption per adult equivalent	.23	.20
	(.14)	(.15)
Poverty head count, basic needs line	−.23**	−.20
	(.11)	(.12)
Poverty head count, food poverty line	−.01	−.02
	(.08)	(.10)
Poverty gap, basic needs line	−.03	−.03
	(.04)	(.05)
Poverty gap, food poverty line	−.01	−.01
	(.03)	(.03)
Subjective poverty ladder	−.29	−.37
	(.51)	(.53)

Note. The first column repeats the complete results from panel A in tables 3–6. The second column presents the results when these models are reestimated with ballot loser households who entered multiple ballots given the same weight as the proportion of winning households who entered multiple ballots. Standard errors are in parentheses.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

TABLE A2

HOW NONAPPLICANTS AND NONCOMPLIERS COMPARE TO APPLICANTS

	Unsuccessful Ballot (1)	Nonapplicants (2)	Nonapplicants with Emigrants (3)	Noncompliers (4)	Unsuccessful without Emigrants (5)	Unsuccessful with Emigrants (6)	*t*-test (2) vs. (1) *p*-value	*t*-test (3) vs. (1) *p*-value	*t*-test (4) vs. (1) *p*-value	*t*-test (6) vs. (5) *p*-value
Household size	7.7	6.8	7.0	7.8	8.1	7.5	.013	.038	.818	.320
Number of adults 18–45	3.2	2.4	2.4	3.3	3.0	3.3	.000	.000	.796	.409
Number of children <18	3.3	3.1	3.0	3.6	4.1	3.0	.337	.243	.634	.018
Number of adults >45	1.2	1.4	1.5	1.0	1.0	1.2	.046	.003	.380	.215
Proportion of 18–45 female	.45	.50	.50	.47	.51	.43	.112	.152	.652	.090
Mean age of adults	29.7	31.0	30.9	30.7	29.9	29.6	.051	.085	.366	.764
Mean years of education 18–45	14.0	12.9	12.9	14.4	14.3	13.9	.000	.000	.328	.311
Index of household wealth	.59	−.35	.02	1.40	.97	.44	.000	.017	.079	.227
Located in Apia	.26	.16	.17	.24	.29	.24	.011	.038	.870	.639
Located in Savai'i	.16	.33	.30	.10	.20	.14	.000	.004	.467	.411
Sample size	121	417	254	29	35	86				

Note. All unsuccessful households are included here including those that would all move if they won the ballot. Nonapplicant emigrant households are defined as any that receive international remittances from former family members.

References

Abdulkadiroglu, Atila, Joshua Angrist, Susan Dynarski, Thomas Kane, and Parag Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics* 126, no. 2:699–748.

Acosta, Pablo, Pablo Fajnzylber, and Humberto Lopez. 2007. "The Impact of Remittances on Poverty and Human Capital: Evidence from Latin American Household Surveys." In *International Migration, Economic Development and Policy*, ed. C. Özden and M. Schiff. Washington, DC: World Bank and Macmillan.

Adams, Richard. 1998. "Remittances, Investment, and Rural Asset Accumulation in Pakistan." *Economic Development and Cultural Change* 47, no. 1:155–73.

Adams, Richard, Alfredo Cuecuecha, and John Page. 2008. "Remittances, Consumption and Investment in Ghana." World Bank Policy Research Working Paper no. 4515, World Bank, Washington, DC.

Akee, Randall. 2010. "Who Leaves and Who Returns? Deciphering Immigrant Self-Selection from a Developing Country." *Economic Development and Cultural Change* 58, no. 2:323–44.

Angrist, Joshua. 2004. "Treatment Effect Heterogeneity in Theory and Practice." *Economic Journal* 502: C52–C83.

Barham, Bradford, and Stephen Boucher. 1998. "Migration, Remittances and Inequality: Estimating the Net Effect of Migration on Income Distribution." *Journal of Development Economics* 55, no. 2:307–31.

Borjas, George, and Bernt Bratsberg. 1996. "Who Leaves? The Outmigration of the Foreign-born." *Review of Economics and Statistics* 78, no. 1:165–76.

Brown, Richard, and Gareth Leeves. 2011. "Comparative Effects of Migrants' Remittances on Composition of Recipient Household Income in Two Small Island Economies." *Applied Economics* 43, no. 27:3965–76.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48, no. 2:424–55.

Dustmann, Christian, and Yoram Weiss. 2007. "Return Migration: Theory and Empirical Evidence from the U.K." *British Journal of Industrial Relations* 45, no. 2:236–56.

Esquivel, Gerardo, and Alejandra Huerta-Pineda. 2007. "Remittances and Poverty in Mexico: A Propensity Score Matching Approach." *Integration and Trade* 27:47–74.

Gibson, John, and David McKenzie. 2011. "The Microeconomic Determinants of Emigration and Return Migration of the Best and Brightest: Evidence from the Pacific." *Journal of Development Economics* 95, no. 1:18–29.

Gibson, John, David McKenzie, and Steven Stillman. 2011. "The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program." *Review of Economics and Statistics* 93, no. 4:1297–1318.

Gundel, Sebastian, and Heiko Peters. 2008. "What Determines the Duration of Stay of Immigrants in Germany? Evidence from a Longitudinal Duration Analysis." *International Journal of Social Economics* 35, no. 11:769–82.

Imbens, Guido. 2010. "Better LATE than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature* 48, no. 2:399–423.

McKenzie, David. 2005. "Beyond Remittances: The Effects of Migration on Mexican Households." In *International Migration, Remittances and the Brain Drain*, ed. Caglar Özden and Maurice Schiff. Washington, DC: World Bank and Macmillan.

McKenzie, David, John Gibson, and Steven Stillman. 2007. "Moving to Opportunity, Leaving Behind What? Evaluating the Initial Effects of a Migration Policy on Incomes and Poverty in Source Areas." *New Zealand Economic Papers* 41, no. 2: 197–223.

———. 2010. "How Important Is Selection? Experimental vs. Non-experimental Measures of the Income Gains from Migration." *Journal of the European Economic Association* 8, no. 4:913–45.

McKenzie, David, and Hillel Rapoport. 2007. "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico." *Journal of Development Economics* 84, no. 1:1–24.

Rubalcava, Luis, Graciela Teruel, Duncan Thomas, and Noreen Goldman. 2008. "The Healthy Migrant Effect: New Findings from the Mexican Family Life Survey." *American Journal of Public Health* 98, no. 1:78–84.

Vaai, Kolone. 2007. "Samoa Economic Update, 2007." *Pacific Economic Bulletin* 22, no. 3:1–14.

Woodruff, Christopher, and Rene Zenteno. 2007. "Migrant Networks and Microenterprises in Mexico." *Journal of Development Economics* 82, no. 2:509–28.

Yang, Dean. 2008. "International Migration, Remittances, and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *Economic Journal* 118, no. 528:591–63.