Measuring Migration from the Republic of Georgia: A Comparison of Three National Surveys[1]

******DRAFT – Please do not cite******

Abstract: Large-scale out-migration is a fairly recent phenomenon in the Republic of Georgia, and social and political turmoil, as well as the diverse nature of migration flows from the country, have made measurement of migration particularly difficult. Three recent national surveys have made it possible to explore patterns of international migration from Georgia at the national level: the Caucasus Research Resource Centers Data Initiative (2007), the Development on the Move survey (2008), and the GeoStat migration survey (2008). However, there are notable differences in the estimates of the prevalence of international migration, as well as in the demographic characteristics of migrants, produced by these three surveys. My paper compares the surveys and assesses their sampling strategies and approaches to measuring migration. I argue that issues of sampling design, representation of ethnic minorities, and differing conceptualizations of "migrant" and "household" across the three surveys explains the differences in results.

Introduction

Migration is one of the more challenging demographic phenomena to measure, due to its often temporary and circular nature, as well as to migrants' desire to avoid detection and enumeration. Among the possible strategies for measuring and studying migration, one that is becoming increasingly common involves household surveys in migrant-sending countries. Several such surveys have recently been done in the Republic of Georgia, where rapidly changing migration patterns present special challenges to measuring migration. My paper will describe three recent national surveys done in Georgia, comparing the approaches used to measure migration and assessing the sampling strategies. The three surveys produce fairly similar estimates of the overall prevalence of migration, but different estimates of the geographic origins of migrants within Georgia

[1] Support for this research was provided by an NSF SBE Doctoral Dissertation Improvement Grant and a Title VIII Combined Research and Language Training grant from American Councils for International Education (ACTR/ACCELS). I would like to thank Irina Badurashvili, Hans Gutbrod, Paata Shavishvili, Karine Torosyan, and Tinatin Zurabishvili for providing access to and assistance with data sets. I also thank Cynthia Buckley, Dustin Brown, Andrea Henderson, and Catherine McNamee for their helpful comments and suggestions.

and of the demographic characteristics of migrants. To a large extent, these discrepancies can be explained by the inherent difficulties of sampling a rare and geographically clustered population, but these three surveys also face specific problems related to the representation of ethnic minority populations and the definition of the concepts "household" and "migrant".

There is a long history of using census and survey data in migrant-receiving countries to look at immigrant populations (Edmonston and Michalowsky 2004). More recently, a growing number of household surveys in migrant-sending countries have been used capture out-migration. In such surveys, household members are asked to report their own migration histories, as well as to provide proxy information on any household member who is a current migrant. Using household surveys to measure out-migration has two key advantages. First, emigration, particularly undocumented migration, is not easily captured by official statistics (Massey and Capoferro 2004). Second, surveys can collect a wide variety of socioeconomic and demographic data about both migrants and non-migrants, allowing for comparison between the two population groups. Nevertheless, measuring migration based on household surveys also has a number of drawbacks and challenges.

The first challenge is the difficulty of creating a representative sample of migrants. For any sub-population that totals less than 10 percent of the total population, specialized sampling methods, are typically required to ensure a representative sample of the sub-population (Kish 1987). Most countries have annual emigration rates below 1 percent, making migrants, particularly recent migrants, a rare element in national populations (UN Statistical Division and Bilsborrow 2009). Stratification and oversampling strategies are often very useful in capturing rare populations. But stratification by migrant status is very

difficult, particularly in countries like Georgia where there are no reliable estimates of the prevalence of migration in the population at large, or the geographic distribution of migrants.

Geographic concentration of migrants presents a specific challenge for the cluster sampling strategies used by most household surveys. Migration is heavily driven by the existence of migration-specific social capital, which is unevenly distributed across households and across communities (Palloni, Massey and Ceballos 2001). Migration is also a cumulative process, with initial migration from a certain community spurring further migration, until a saturation level is reached (Massey 1990). While migration can and does spread from one sending community to nearby communities, the prevalence of migration can vary widely from community to community, even within a small geographic area. This is especially true when large-scale migration outflows are a recent development, as they are in Georgia. When migrants are heavily clustered in certain primary sampling units (PSUs) in a survey, sample variance will be high and population estimates will suffer from large confidence intervals.

The second challenge of measuring migration using a sending-country based survey is the fact that many of the desired respondents are absent from the country and therefore unable to participate in the survey. For respondents who are returned migrants, surveys can record migration histories, but current migrants can only be captured if there is someone remaining behind who is able and willing to provide information about the migrant. When a whole household migrates, no one in that household is likely to be captured by a sending-country survey. Even when a migrant's family members remain behind, they may be unwilling to report on their migrant relative, due to social stigma or

fear of legal consequences, or they may have little about the migrant's background or current whereabouts. Alternately, respondents in the sending country may simply no longer think of the migrant as a member of the household; surveys commonly ask about the current location of people who "usually" live in the household, although it is likely that this type of phrasing leads to an under-count of long-term migrants (Massey and Zenteno 2000).

Despite these important limitations, household surveys—even those that do not specifically attempt to identify a representative sample of migrants—can be effectively used in migration research (Bohra and Massey 2009; Curran et al. 2005; Frank and Hummer 2002; White and Buckley 2011).

Measuring migration in Georgia

Georgia is a former Soviet republic with a population of approximately 4.6 million, located in the southern Caucasus region, between Russia and Turkey. During the Soviet period, there was little international migration from Georgia, although there was some internal migration, mostly to Russia. Since becoming independent in 1991, Georgia has faced severe economic collapse, inflation, energy shortages, political turmoil, and civil war. Currently two *de jure* regions of the country—South Ossetia (known in Georgia as Tskhinvali region) and Abkhazia—remain outside of the control of the central government, and wars in these regions have displaced tens of thousands of people. Since 2000, Georgia has attained a degree of stability and achieved some economic development, but local employment opportunities remain very limited. These conditions provide strong push factors for labor migration, leading to high levels of out-migration from Georgia throughout the post-Soviet period.

Because large-scale out-migration is a fairly recent phenomenon, because of the social and political turmoil of the country, and because Georgian migrants travel to a wide variety of destination countries, measuring migration from Georgia has been particularly difficult. Although an ongoing IOM project is attempting to improve the quality of official migration statistics (IOM 2009), currently the only indicator of migration published by the Georgian national statistical office (GeoStat) is annual net migration. Information about Georgian migrants from statistics collected in their destination countries is also very limited. Due to the very small number of Georgians in any single destination country, and the fact that many are unauthorized migrants, Georgians appear in very small numbers in censuses and sample surveys done in destination countries, if they appear at all. In the United States, for example, immigrants from Georgia are often counted together with other immigrants from the former Soviet Union, making them indistinguishable from the larger Russian- and Ukrainian-origin populations.

Household surveys done in Georgia therefore provide the best opportunity to understand migration patterns from Georgia, and study the characteristics of Georgian migrants, but until recently, none were available. However, several previous small-scale studies point to a number of interesting trends (Badurashvili 2004; Dershem and Khoperia 2004; IOM 2003; Zurabishvili, Tavberidze and Zurabishvili 2009). Labor migration from Georgia appears to be undergoing a process of feminization, with men dominating migration flows in the 1990s and women playing an increasing role since 2000. This is occurring at the same time as a dramatic shift in migrant destinations away from Russia, which was the primary destination in the 1990s, and towards Turkey,

Greece, Western Europe, and North America. While men still dominate migrant flows to Russia, women comprise the majority of migrants to other destinations. This combination of feminization and destination shift is unusual, and has received little attention in the migration literature, making Georgia a fascinating place to study migration patterns.

Three new data sets

Starting in 2007, several national household surveys have included migration questions, making it possible for the first time to explore patterns of out-migration from Georgia at the national level. There are three data sets that can be used to explore migration patterns from Georgia:

Caucasus Research Resource Center (CRRC) Data Initiative (DI), 2007. The DI has been conduced annually since 2004 in the three countries of the South Caucasus (Armenia, Azerbaijan, and Georgia). The sampling methods and survey questions have varied substantially from year to year, and only the 2007 survey includes migration questions. The 2007 DI in Georgia used a multi-stage cluster sample of households, using the Census Instructor Areas from Georgia's 2002 census as PSUs. Because maps of the Census Instructor Areas are not publicly available, the PSUs were reconstructed by CRRC. The sample was stratified by residence in a rural area, an urban area (but not the capital city), or the capital city, as well as by residence in the each of four geographic quadrants (north-east, north-west, south-east, south-west). This resulted in a total of nine strata, with a total sample of 3,392 households. Sampling weights were calculated to account for the varying sizes of the PSUs and were adjusted for non-response (the response rate was 83%) (CRRC 2007).

The 2007 DI questionnaire, which was available in Armenian, Azeri, Georgian, and Russian, included information on whether each household member had ever lived abroad, as well as a special supplement for members of the household who were currently located abroad. Interviewers were instructed to ask respondents to distinguish between "people presently living with you most of the time, regardless of legal place of residence, and who share income and expenses", who were considered present household members, and people "whom you consider at the present time to be members of your household, who contributes to support your household or whom the household contributes to support, but is not currently living with the rest of the household and who has been gone for three months or more", who are considered absent migrants. Limited demographic information on absent migrants was provided by proxy by another household member. The survey also includes household socioeconomic characteristics, as well as a large number of questions on the values and political beliefs of a sample household member.

GeoStat Migration Survey, 2008. Georgia's national statistical office, GeoStat, conducted its own migration survey in 2008, although the data has not been widely distributed, even within Georgia. Unfortunately the available documentation for this survey is very limited. The GeoStat survey is a multi-stage cluster sample, using the 2002 Census Instructor Areas as PSUs, and stratified to represent Tbilisi and 10 other regions of Georgia (GeoStat 2009a). The final sample includes 5,450 households. The GeoStat survey defined absent migrants as household members who were currently living in another country and had been gone, or intended to be gone, for at least six months. Information on all household members, including absent migrants, was provided by a single household respondent. The survey includes a very small number of variables,

covering only basic demographic characteristics, household income, basic migration history and reasons for migration. The questionnaire was available in Georgian only, and the response rate was 78 percent.

Development on the Move: Measuring and Optimizing Migration's Economic and Social Impacts in Georgia (DOTM), 2008. This survey was sponsored by the Global Development Network (India) and the Institute for Public Policy Research (UK), and the fieldwork was conducted by CRRC. It is the only survey of the three to stratify households by migration status and oversample migrant households.. The survey was based on Georgia's electoral precincts, which were used as PSUs. A total of 42 PSUs were randomly selected, and stratified by location in rural areas, urban areas, and the capital city. Some PSUs were excluded due to a high proportion of non-Georgian speakers (3.9% of all PSUs), and others were excluded because they were inaccessible while fieldwork was being conducted (fall 2008) due to the conflict between Georgia and Russia (3.8%). Interviewers screened every household in the 42 selected PSUs to determine if the household included any current or former migrants. Based on that original screening, approximately 500 households without migrants, 500 with current migrants, and 500 with returned migrants were selected. The final sample size is 1484 households. Weights were calculated based on the proportion of migrants in the original screening in order to develop estimates of the prevalence of migration at the national level (Tchaidze and Torosyan 2010).

The DOTM questionnaire, which was available in Georgian only, includes socioeconomic and demographic data, and a wide variety of migration-specific questions. Absent migrants were defined as members of the household currently residing in another

country, who had been gone for at least three months. Information on absent migrants was collected by proxy from another household member. The response rate for the DOTM was the lowest of the three surveys, at 75 percent.

Table 1 summarizes the main characteristics of all three surveys. There are four key differences that should be mentioned. First, while the DOTM and GeoStat surveys were both conducted in late 2008, the DI was conducted a year earlier – prior to the global economic crisis of 2008 and prior to the August 2008 war between Russia and Georgia. Second, there are substantial differences in sampling across the three surveys. None used the same set of PSUs; although both the DI and the GeoStat survey were based on 2002 census instructor areas, the boundaries of these areas had to be reconstructed from partial data when the DI sampling strategy was developed. None of the three used the same strata, with the DI and the DOTM focusing more on rural-urban stratification and the GeoStat focusing on administrative regions. The DOTM is particularly different from the other two in that it over-sampled migrant households and because it excluded areas where a high percentage of ethnic minorities live. Third, the GeoStat survey is notable for its lack of sampling weights, which the other two surveys include. Finally, there are key differences in the definition of "migrant". The GeoStat survey requires that an individual be gone for six months to qualify as a migrant, while the DI and DOTM require only three months. Additionally, the DI questionnaire specifies that someone who "contributes to the support of the household or the household contributes to support" constitutes a member of the household; migrants are therefore people who do not reside with the household, but are financially linked to it. The other two surveys leave open the question of what constitutes a household member.

In the following section, I will describe the similarities and differences in the results of the three surveys, focusing on the following areas: 1) estimates of the overall prevalence of international migration at the national level; 2) geographic origin of international migrants within Georgia; 3) basic demographic characteristics of migrants (age, sex, education level, marital status, and year of departure). All analyses use weighted data (except in the case of the GeoStat survey, where no weights are available) and adjust for clustering and stratification, using svy estimates in Stata IC 10.1.

Descriptive statistics

The three surveys produce similar estimates of the overall prevalence of international migration (see Table 2), although the DI estimates are slightly smaller than those produced by the other two surveys. According to the DI, 7.47% of all households include an absent migrant; the GeoStat and DOTM estimates are slightly larger, at 9.93% and 9.62%, respectively. The difference between the DI and the GeoStat is statistically significant ($p<.05$), but the difference between the DI and the DOTM is not. Similarly, the DI estimates that 3.26% of all adults are absent migrants, while the GeoStat estimate is 4.84% and the DOTM is 3.92%. In this case, all three estimates are statistically identical.

The fact that three such different surveys produce similar estimates of the prevalence of migration is heartening. However, when you compare the characteristics of the migrants identified in the three surveys, notable differences start to appear. Table 3 shows the distribution of absent migrants across geographic regions of Georgia. The DOTM does not provide data on region, and a comparison between the GeoStat and DI data is complicated by the fact that each defines "region" differently. The GeoStat survey's

"region" variable includes 10 regions: the capital city of Tbilisi and 9 other administrative regions of Georgia. The DI region variable includes only five: Tbilisi and four geographic "quadrants" that each encompass two or three administrative regions. The confidence intervals around estimates at the quadrant level are very large, so there are no differences in the regional distribution of migrants that are significant at the .05 level. However, the point estimates are quite different in several cases. Both surveys find about 20 percent of all migrants originating in Tbilisi, and around 10 percent originating in the South-East (regions of Kvemo Kartli and Samtskhe-Javakheti). However, the DI finds a higher proportion of migrants originating in the three Northeastern regions (36% versus 24%), while the GeoStat survey identifies a much larger proportion of migrants in the Northwestern regions of Samegrelo and Imereti (32% versus 15%).

Examining the demographic characteristics of migrants (Table 4) produces even more discrepancies. The only demographic characteristics available in all three surveys are age, sex, and year of departure (for the most recent international trip). Judging the similarity or difference between the three surveys is hampered by very large confidence intervals, particularly in the DOTM data. However, there are some notable differences in point estimates, even when the surveys do not significantly differ from each other. All three surveys produce very similar estimates of mean age of migrants (around 38 years). The DI and DOTM estimate that a similar proportion of migrants are women (39% and 35%), while the DOTM estimate is notably higher (44%), although not different from either of the other two at the .05 level. Estimates of departure cohorts of migrants are also fairly similar, with all three surveys showing a clear trend towards increased migration in the most recent cohorts. There is one notable difference, however: the DOTM includes a

much smaller proportion of migrants leaving between 1994 and 1998 than do the other two surveys.

The GeoStat and DOTM surveys also include information on educational attainment and marital status for absent migrants, and these variables show dramatic differences. The DOTM shows a much more highly educated sample than does the GeoStat. In the GeoStat data, half of all migrants hold a high school degree or less, with an additional 15 percent having some postsecondary education and 35 percent holding a college degree. In the DOTM data, significantly fewer migrants – around 30 percent – have a high school degree or less, with 27 percent having some postsecondary education and 42 percent having a college degree. Marriage rate are also different, although not significantly so; over 70 percent of migrants in the DOTM sample are married, compared with 65 percent in the GeoStat sample.

The descriptive statistics highlight three areas of concern. First is the small difference in estimates of overall prevalence of migration between the DI and the other two surveys. While the GeoStat and DOTM surveys find absent migrants in just under 10 percent of households in Georgia, the DI finds migrants in about 7.5% of households. Second are the differences in the geographic distribution of migrants in the DI and GeoStat surveys, particularly their very different estimations of migration in the Northwestern part of Georgia. Third are the different estimates of the demographic characteristics of migrants in the GeoStat and DOTM surveys. While both surveys produce very similar estimates of the prevalence of migration, the DOTM indicates that migrants are more likely to be male, more likely to be married, and more likely to have postsecondary education than does the GeoStat survey.

Explaining the discrepancies

One obvious potential explanation for the discrepancy between the DI and the other two surveys is the time that they were conducted: the DI in late 2007, the others in late 2008. If there happened to be an unusually high level of out-migration during 2008, the GeoStat and DOTM surveys would report a higher level of migration than the DI. Both the GeoStat and DOTM surveys enumerated returned migrants as well as absent migrants, allowing us to identify the level of migration that these surveys would have identified had they been taken in late 2008 instead of late 2008. Both surveys identify a fairly large proportion of absent migrants who left in 2008, but a nearly equal number of returned migrants who were absent in 2007 and returned in 2008, so the overall prevalence of migration remains nearly the same for both years. Based on this information, I will treat the one-year difference between the DI and the other two surveys as irrelevant, and explore three other potential sources of difference between the two surveys: sampling issues, representation of ethnic minorities, and differences in definitions of the terms "migrant" and "household".

Sampling issues

As may be expected, in all three surveys, there is substantial clustering of migrants within PSUs, and this is especially true in the GeoStat data. In the GeoStat survey, 51 percent of migrants are identified in only the top 10 percent of PSUs, while in the DI 34 percent of migrants are identified in 10 percent of PSUs, and 27 percent in the DOTM. This high degree of clustering creates very large sampling variances, and very large confidence intervals. The problem is further complicated by the substantial number of PSUs in the

DI and GeoStat samples that contain no migrants: 21 percent of PSUs in the DI and 28 percent in GeoStat. The DOTM, with its deliberate oversample of migrants, largely avoids this problem; only three of its 42 PSUs include no absent migrants. These differences are reflected in the design effects for the estimates of migration: 4.19 for the GeoStat survey, versus 3.10 for the DI and 2.91 for the DOTM. The sampling design DOTM clearly has advantages over the other two surveys, in terms of estimating migration, but these design advantages are offset by the small sample size of the DOTM, and the confidence intervals around DOTM survey estimates are similar to those produced by the other surveys.

The extreme sensitivity of the DOTM estimates to the inclusion of sampling weights is also notable. The GeoStat survey does not have weights, and the point estimates provided by the DI change little whether the sample weights are included or not. The DOTM estimates, however, are highly sensitive to the weighting. Given the over-sample of migrant households, the prevalence estimates are of course sensitive to weighting, but the estimates of the demographic characteristics of migrants are also sensitive. Table 5 shows point estimates of the demographic characteristics of migrants in the DOTM survey, comparing weighted and unweighted estimates. Compared to the weighted estimates, the unweighted estimates show about 5 percent more women among migrants, as well as a greater concentration of migrants at lower education levels, and a lower proportion of migrants married. In short, without the weights, the migrant sample in the DOTM looks more like the GeoStat and DI samples.

Under-representation of ethnic minorities

The three surveys may differ substantially in their representation of Georgia's ethnic minority populations. Georgia has substantial Armenian, Azeri, and Russian populations, as well as other minority groups. Many ethnic minorities, particularly those living in rural areas, do not speak Georgian (Tchaidze and Torosyan 2010). The DI was the only survey of the three to provide questionnaires in languages other than Georgian. The GeoStat survey used only Georgian language questionnaires, but did not specifically exclude minorities. The DOTM excluded from their sample the nearly 4 percent of PSUs where over 50 percent of the population are ethnic minorities, and therefore deliberately under-represents Georgia's minority populations. If ethnic minorities are less prone to migration than ethnic Georgians, this may explain the slightly smaller estimates of migration in the DI. If ethnic minorities differ from Georgians in terms of education level or other demographic characteristics, that might explain the differences between the GeoStat and DOTM surveys in terms of demographic characteristics of migrants.

The DI does not record ethnicity for all household members, but it does record the ethnicity of one randomly selected adult in each household, and the language of interview (Georgian, Armenian, Azeri, or Russian). In the DI sample, just under 84 percent of sampled adults identify as Georgian, 8 percent as Armenian, 3 percent as Azeri, 2 percent as Russian, and 3 percent as other (see Table 6). Based on the ethnic breakdown of the population in Georgia's 2002 census, the sample slightly over-represents Armenians (5.7% of the population in the census) and under-represents Azeris (6.5% in the census) (GeoStat 2009b). Interestingly, households where someone identifies as Armenian are more likely to include an absent migrant than households where the respondent identifies

as any other nationality, although this difference is not significant.[2] This over-representation of Armenians therefore cannot explain why the DI produces lower estimates of migration than the other two surveys.

The GeoStat survey, which records the ethnic identification of every member of the household, does not seem to greatly under-represent ethnic minorities in comparison with the DI (see Table 5). Unlike the DI sample, in which only Armenians and "Others" migrate at a higher rate than Georgians, in the GeoStat sample, members of all ethnic groups are more likely to migrate than are Georgians. The GeoStat survey identifies a higher propensity to migration among all ethnic groups than does the DI, so differential representation of ethnic minorities cannot explain the differences between these two surveys.

The DOTM survey did not record ethnicity, but did record language abilities of household members (without distinguishing between native language and second languages). Although there is not enough information to judge how well ethnic minorities are represented, the fact that only 2.78% of adults in the sample (most of whom reside in Tbilisi) speak Armenian, and only 1.28% speak Azeri, is suggestive of serious under-representation of rural minority communities, who may migrate at higher rates than Georgians. This may explain the slightly higher prevalence of migration among adults identified in the GeoStat survey, in comparison with the DOTM survey. The under-representation of rural minority populations in the DOTM sample may also partially explain the substantial differences in demographic characteristics of migrants between the two surveys. If non-Georgians are excluded from the GeoStat sample, the educational differences between migrants in the GeoStat and DOTM samples become statistically

[2] Results based on language of interview are not shown, but they are similar

insignificant at the .05 level. However, the GeoStat survey still finds that a higher proportion of migrants are women, and that fewer migrants are married, than does the DOTM survey.

Although the DOTM sample clearly under-represents ethnic minorities, in comparison to both the GeoStat and the DI surveys, this seems to play little role in explaining the differences between the three surveys. Differences in both prevalence estimates and the demographic profile of migrants between the GeoStat and DOTM may be partly explained by ethnic differences, but ethnic differences cannot explain why the DI produces the lowest overall estimates of migration prevalence.

Definitional issues

One key definitional issue in migration surveys is exactly how long a person needs to be absent in order to qualify as a "migrant." In these surveys, the GeoStat uses the most restrictive definition, requiring an absence of at least six months, while the other two surveys require only three. The obvious implication is that the GeoStat survey should produce the smallest estimates of migration, because it excludes a certain number of migrants that would be included in the other two surveys. This, however, is not at all the case; the GeoStat survey provides the highest estimates of migration. In all three surveys, from 10-15 percent of absent migrants left within the past six months; more recent migrants seem to be included in the GeoStat, probably because the respondent believes they intend to be gone for at least six months, even if they have not yet been gone for so long.

A more serious definitional issue in these three surveys arises from the fact that information on current migrants must be collected by proxy. A straightforward question

along the lines of: "Are any members of your household currently living abroad?" may not be simple for a respondent to answer. A respondent may have many relatives living abroad, but which should be considered "household members"? Household is a somewhat ambiguous term, and the more so in Georgia, where households are often multi-generational and where there is even no good native translation of the term "household." Small variations in how the household roster was defined and recorded could therefore have important consequences for which migrants were captured and which were not.

As Table 7 shows, there is a dramatic difference between the DI and the other two surveys in who the migrants are in the household. Over 75 percent of all migrants identified by the DI are children (or children-in-law) of the household head, compared to under 60 percent in the GeoStat and DOTM surveys. Only 13 percent of migrants in the DI are senior members of their households, compared with 25 percent in the GeoStat and 30 percent in the DOTM, indicating that the DI might for some reason under-represent this group of migrants. The DI also seems to have a greater representation towards more recent migrants: 30 percent of absent migrants in the DI sample have been gone fewer than two years, compared with 26 percent in the GeoStat data and only 20 percent in the DOTM.

Why the DI might under-represent migration in the older generation of a household is very difficult to judge, especially because documentation on how any of the three surveys instructed their interviewers to record household members is limited at best.[3] However, it is worth noting that the DI, unlike the other two surveys, included a specific definition of "household" in its questionnaire. The DI definition of household, which revolved around

[3] I will be making a trip to Georgia in the near future, and will attempt to learn more about how interviewers were trained.

shared financial support, may have led some respondents to decide against listing migrants whom they might otherwise have considered to be household members. Also worth noting is the fact that the DI was not specifically a migration survey, while the other two were – and would have been introduced to the respondents as such. Knowing that the goal of the survey was to count migrants, respondents may have made an effort to list migrants in their households, which they may not have done for a survey where counting migrants was not presented as the goal.

Conclusions

By comparing the results of three national household surveys, we can obtain a reasonable estimate of the prevalence of migration from Georgia: in the late 2000s, approximately 10 percent of households in Georgia included a member who was currently living abroad, equaling roughly 3-5 percent of all adults living abroad. The fact that three surveys, using different sampling frames, different strata, different questionnaires, and all with their own limitations, generated such similar estimates, gives confidence those estimates. Discrepancies among the three surveys in terms of regional origin and demographic characteristics of migrants (particularly education level and sex breakdown) are more difficult to reconcile.

I identify four main problems with the three surveys that may explain these discrepancies. First, the sampling weights in the DOTM may introduce error into the estimates. The DOTM weights are designed to correct for the oversample of migrant households and produce accurate national estimates of the prevalence of migration, but they do not take into account the sex or other characteristics of migrants. Because migration is selective on characteristics such as sex, age, and education, and because the

nature of this selectivity varies geographically, not including this information in the weights may limit the accuracy of the weighted estimates. Second, the DOTM seriously under-represents rural ethnic minorities, which may lead to a sample of migrants who are more highly educated than the other two surveys. Third, the question of who constitutes a household member and who constitutes a migrant seems little thought-out in all three surveys, and particularly in the GeoStat and DOTM surveys. Finally, the very large sampling error in around most estimates of migration greatly limits the utility of demographic data about migrants, even in the relatively large GeoStat sample. The main reason for large sampling error in the GeoStat and DI samples appears to be clustering of migrants within PSUs; clustering is less of a problem in the DOTM sample, but the DOTM sample size is so small that confidence intervals are still quite large.

To some extent, these types of problems are inevitable. Capturing a rare and rapidly changing sub-population with a household survey is inherently challenging. In the United States, research on the characteristics of the U.S. Hispanic population is hampered by the fact that each different survey appears to capture a slightly different Hispanic sub-population, leading different surveys to draw sometimes contradictory conclusions. This problem exists even in large, nationally representative samples and is attributed to the high mobility and uneven geographic dispersion of the Hispanic population (McConnell and LeClere 2002; Perl, Greely and Gray 2006). Migrants are also a mobile and unevenly dispersed sub-population, and this type of problem is all the greater in a small country like Georgia. Just as every U.S. survey seems to capture a slightly different group of Hispanics, each of these three Georgian surveys seems to capture a different group of migrants.

Nevertheless, there are important lessons here to improve future migration research in Georgia and beyond. First, it is important in Georgia to consider that different ethnic groups experience different migration patterns. In order to fully understand migration among minority groups living in specific geographic areas, a deliberate attempt to sample these groups will be required. Second, stratification of samples by household migration status is a very complicated process, as the DOTM survey demonstrates, but it has a great deal of potential value. The geographic concentration of migrants in traditional cluster samples is problematic because it leads to large design effects and very high sampling variance, creating large confidence intervals around all estimates related to migration. Stratified samples have the potential to generate more precise and therefore more useful estimates.

Finally, designers of household surveys who wish to measure migration should take into careful consideration who, exactly, they want to record as "migrants" and how their respondents are likely to understand concepts of "household members" and "migrants." The Mexican migration project gets around the problem of whether or not a migrant should be considered a "household member" by asking household heads about the current location of all their adult children (Massey and Zenteno 2000). Regardless of what approach is taken, information on how the household roster is obtained should be included in survey documentation. Without this information, it is difficult to asses who, exactly, is being counted in surveys that try to measure migration.

Table 1. Characteristics of the DI, DOTM, and GeoStat surveys

	DI	DOTM	GeoStat
Year	2007	2008	2008
Sampling frame	Civilian, non-institutionalized population of Georgia, minus Abkhazia and South Ossetia	Civilian, non-insitutionalized population of ethnic Georgian-majoirty areas of Georgia, minus Abkhazia, South Ossetia, and surrounding territories occupied by the Russian army	Civilian, non-institionalized population of Georgia, minus Abkhazia and South Ossetia
Sample size	3,392 households	1,484 households	5,454 households
Stage 1 frame	2002 Census instructor areas (reconstructed)	Electoral districts for 2008 parliamentary elections	2002 Census instructor areas
PSUs	81	42	303
Strata	9 (Northwest-urban, Northwest-rural, Northeast-urban, Northeast-rural, Southwest-urban, Southwest-rural, Southeast-urban, Southeast-rural, Capital)	3 strata (urban, rural, capital) and 3 sub-strata (households with no migrants, households with current migrants, households with returned migrants)	10 (Tbilisi, Achara, Guria, Imereti, Kakheti, Mtskheta-Mtianeti, Samegrelo and Svaneti, Samtskhe-Javakheti, Kvemo Kartli, Shida Kartli)
Response Rate	83%	75%	78%
Post-survey	weights, clusters	weights, clusters	clusters

Table 2. Prevalence of migration in Georgia

	DI	GeoStat	DOTM
Total households	3,291	4,454	1,482
% of households including an absent migrant	7.47 (6.01-8.93)	9.93 (8.67-11.20)	9.62 (7.11-12.13)
Total adults	9,807	16,213	4,327
% of adults who are absent migrants	3.26 (2.64-3.88)	4.84 (4.20-5.57)	3.92 (2.93-4.91)

Table 3. Regional distribution of migrants

		Percentage of all absent migrants	
		GeoStat	DI
Tbilisi		21.11 (16.51-26.58)	20.21 (13.74-28.75)
North-East	Mtskheta-Mtianeti	10.3 (5.79-17.66)	36.34 (27.43-46.29)
	Kakheti	6.95 (4.04-11.70)	
	Shida Kartli	6.31 (3.83-10.21)	
North-West	Samegrelo and Zemo Svaneti	10.42 (6.38-16.59)	15.05 (9.50-23.01)
	Imereti	21.24 (15.47-28.43)	
South-East	Kvemo Kartli	6.05 (3.92-9.22)	9.07 (5.85-13.81)
	Samtskhe-Javakheti	4.12 (2.14-7.80)	
South-West	Achara	9.27 (6.16-13.72)	19.31 (12.46-28.70)
	Guria	4.25 (2.78-6.45)	

Table 4. Selected demographic characteristics of migrants

	DI	GeoStat	DOTM
% female	38.64 (32.27-45.01)	44.14 (40.58-47.71)	35.39 (27.29-43.49)
mean age	38.37 (37.03-39.71)	38.57 (37.59-39.55)	37.91 (35.98-39.83)
Year of departure			
before 1989	.88 (.28-2.77)	2.06 (1.24-3.02)	3.78 (.77-16.48)
1989-1993	5.14 (2.69-9.61)	6.06 (4.09-8.90)	9.66 (2.66-24.88)
1994-1998	20.72 (13.99-29.58)	16.65 (12.54-21.76)	8.51 (6.24-11.52)
1999-2003	23.56 (18.07-30.10)	24.26 (.19.92-29.19)	26.46 (19.19-35.52)
2004-2008	49.69 (39.02-60.39)	50.97 (45.29-56.62)	51.50 (40.36-62.49)
Education			
Less than HS	x	4.52 (3.02-6.72)	2.59 (.70-9.09)
High school	x	44.96 (39.74-50.29)	27.44 (20.15-36.17)
Some college	x	15.12 (12.31-18.41)	27.48 (19.45-37.29)
College degree	x	35.40 (30.40-40.74)	42.49 (31.72-54.02)
% married	x	64.74 (60.55-68.71)	70.34 (63.87-76.81)
Weighted n	319	786	170

Table 5. Comparison of Weighted and Unweighted Estimates in the DOTM

	Unweighted	Weighted
% female	40.71	35.39
mean age	38.05	37.91
Departure cohort		
before 1989	1.41	3.78
1989-1993	6.83	9.66
1994-1998	13.31	8.51
1999-2003	26.44	26.46
2004-2008	52.01	51.5
Education		
Less than HS	1.75	2.59
High school	28.77	27.44
Some college	33.86	27.48
College degree	35.61	42.49
% married	64.7	70.34
N	592	170

Table 6. Distribution and prevalence of migration by ethnic group

	DI		GeoStat	
	% in sample	% migrants	% in sample	% migrants
Georgian	83.84	2.86	84.24	3.98
Armenian	7.85	6.54	6.15	10.73
Azeri	3.29	2.04	6.08	7.23
Russian	1.92	1.35	1.12	7.29
Other	3.1	4.66	2.41	14.86
Total	100	3.16	100	4.88

Table 7. Percent of migrants, by household category

	DI	DOTM	GeoStat
Household head or spouse of head	13.13	30.23	25.2
Child or child-in-law of household head	75.42	54.73	58.65
Other relative	11.46	14.7	15.91
Total	100	100	100

References

Badurashvili, I. 2004. "Determinants and Consequences of Irregular Migration in a Society under Transition: The Case of Georgia." in *Population Association of America Annual Meeting*. Philadelphia, PA.

Bohra, P.and D.S. Massey. 2009. "Processes of Internal and International Migration from Chitwan, Nepal." *International Migration Review* 43(3):621-651.

CRRC. 2007. "Data Initiative 2007: Unified Sampling Methodology." Tbilisi: Caucasus Research Resource Centers.

Curran, S.R., F. Garip, C.Y. Chung, and K. Tangchonlatip. 2005. "Gendered Migrant Social Capital: Evidence from Thailand." *Social Forces* 84(1):225-255.

Dershem, L.and T. Khoperia. 2004. "The Status of Households in Georgia - 2004." Tbilisi: Save the Children, USAID, and the Institute for Polling and Marketing.

Edmonston, B.and M. Michalowsky. 2004. "International Migration." in *The Methods and Materials of Demography*, edited by J.S. Siegel and D.A. Swanson. Amsterdam: Elsevier Academic Press.

Frank, R.and R.A. Hummer. 2002. "The Other Side of the Paradox: The Risk of Low Birth Weight among Infants of Migrant and Nonmigrant Households within Mexico." *International Migration Review* 36(3):746-765.

GeoStat. 2009a. "migraciis gamokvlevis angarishi." Tbilisi: GeoStat.

—. 2009b. "Statistical Yearbook of Georgia: 2009." Tbilisi: Department of Statistics under the Ministry of Economic Development of Georgia.

IOM. 2003. *Labour Migration from Georgia*. Tbilisi: IOM.

—. 2009. "Capacity Building Programme." Tbilisi: International Organization for Migration.

Kish, L. 1987. *Statistical Design for Research*. New York: John Wiley & Sons.

Massey, D. 1990. "Social Structure, Household Strategies, and Cumulative Causation of Migration." *Population Index* 56:3-26.

Massey, D.S.and C. Capoferro. 2004. "Measuring Undocumented Migration." *International Migration Review* 38(3):1075-1102.

Massey, D.S.and R. Zenteno. 2000. "A Validation of the Ethnosurvey: The Case of Mexico-U.S. Migration." *International Migration Review* 34(3):766-793.

McConnell, E.D.and F.B. LeClere. 2002. "Selection, context, or both? The Engoish fluency of Mexican immigrants in the American Midwest and Southwest." *Population Research and Policy Review* 21:179-204.

Palloni, A., D.S. Massey, and M. Ceballos. 2001. "Social Capital and International Migration: A Test Using Information on Family Networks." *American Journal of Sociology* 106(5):1262-1298.
Perl, P., J.Z. Greely, and M.M. Gray. 2006. "What Proportion of Adult Hispanics Are Catholic? A Review of Survey Data and Methodology." *Journal for the Scientific Study of Religion* 45(3):419-436.
Tchaidze, R.and K. Torosyan. 2010. "Development on the Move: Measuring and Optimising Migration's Economic and Scoial Impacts in Georgia." Tbilisi: International School of Economics of Tbilisi State University.
UN Statistical Divisionand R.E. Bilsborrow. 2009. "Technical Report on the Use of Censues and Surveys to Measure International Migration." New York: United Nations.
White, K.and C.J. Buckley. 2011. "Exposure to International Migration and its Effect on Childbearing in Turkey." *International Migration Review* In press.
Zurabishvili, T., G. Tavberidze, and T. Zurabishvili. 2009. "Tianeti Household Census 2008 & Tianeti Migrants to Greece 2008 Fieldwork Report." Tbilisi: International Organization for Migration.