Refugee or Internally Displaced Person?
To Where Should One Flee?

Will H. Moore and Stephen M. Shellman
Department of Political Science
The Florida State University
USA
will.moore@fsu.edu & http://garnet.acns.fsu.edu/~whmoore/
-and-
sshellman@mindspring.com & http://garnet.acns.fsu.edu/~sms6202/

22 November 2002

—Discussion Draft—
Please do not cite without first checking to see if a revised version is available

Paper presented at the "Forced Migration, Global Security, and Humanitarian Assistance" Conference, Center for Comparative Immigration Studies, University of California, San Diego, 3 December 2002.

An earlier version of this paper was presented at the Annual Meeting of the International Studies Association, New Orleans, 24 – 27 March 2002. We would like to thank Bill Berry, Tom Carsey, Christian Davenport, Kristian Gleditsch, Mark Lubell, Lanny Martin, and participants at the "Identifying Wars: Systematic Conflict Research and Its Utility in Conflict Resolution and Prevention" Conference, and seminars at UNHCR, Florida State, and Rice University for helpful comments on this study. Bela Hovy kindly made the UNHCR's data available, Susanne Schmeidl graciously provided us with her IDP data, and Barbara Harff kindly shared her genocide/ politicide data with us. Brad Ellis provided able research assistance. We alone are responsible for interpretation and any errors.

Abstract

This study begins with the observation that while we have a reasonable understanding of the circumstances that lead countries to produce forced migration flows, scholars have yet to investigate the circumstances the lead some countries to produce a large number of internally displaced persons and relatively few refugees (or asylum-seekers) as opposed to a large number or refugees and relatively few internally displaced persons. The study builds on our previous work on forced migration to explore this question, and focuses on three groups of variables across two settings: violence; socio-econo-political opportunities; and transaction costs, both locally and neighboring countries. We develop hypotheses and then conduct statistical analyses of those hypotheses using data on a global sample of countries covering the years 1970-1995. We briefly discuss the implications of our findings for contingency planning and the debate about 'economic refugees.'

This study explores a heretofore unexamined issue in forced migration. Consider the following. In the spring and summer of 1998 over 40,000 Kosovars abandoned their homes and hid in forests and the mountains, setting up their own makeshift camps, or seeking shelter with relatives, while virtually no Kosovars crossed an international border in seek of refuge from the fighting (CNN 1999, USCR 1999:247-8). The following spring and summer (1999) an estimated 857,000 fled Kosovo, while another 580,000 abandoned their homes, but remained in the province (Reliefweb, 2000). Thus, one year the Kosovo region of the former Yugoslavia produced a far greater number of internally displaced persons (IDPs) than refugees (we define terms below). Yet in the following year the same region produces far greater numbers of refugees than IDPs. What can account for the dramatic shift in the destinations of Kosovars between 1998 and 1999?

This question has gone unexplored in the social science literature on forced migration. Political science has paid limited attention to the topic of forced migration, though there are some prominent exceptions (e.g., Zolberg, Suhrke & Aguayo 1989) and a small literature does exist.[1] However, there is a substantial interdisciplinary literature on forced migration. Yet in that literature—as in the nascent political science research on the topic—the primary focus of causal theory has been on the causes of refugee flows and, to a lesser extent, IDP flows. In this study we take up the task of identifying the characteristics of countries that lead them to produce a high proportion of internally displaced persons relative to refugees, and vice versa.

[1] See Weiner (1995, 1996), Gibney, Apodaca & McCann (1996), Apodaca (1998), and Davenport, Moore & Poe (2002).

Let us define some terms. A forced migrant is one who, owing to a reasonable fear of persecution—by which we mean deprivation of life, liberty, or physical being—has abandoned her/his possessions and/or domicile and relocated either within her/his own country or sought asylum abroad. A refugee is a forced migrant who seeks asylum abroad whereas an internally displaced person (IDP) is a forced migrant who relocates within the borders of her/his country of origin. The research question that we tackle here is what characteristics help identify countries that will produce a large number of refugees relative to IDPs?

While there has been no systematic, large-N research on our question, Susanne Schmeidl (2000:152) reports that "refugees and IDPs flee from similar root causes rather than responding to completely different occurrences." The differences between the size of the two types of migrant populations across different countries of origin, then, are explained by "the conditions that allow (or discourage) exodus across international borders," particularly "internal conflicts in potential countries of asylum…donor fatigue…impassible border areas… [and] a more diverse uprooted population…less equipped to travel long distance" (Schmeidl 1998:33). Schmeidl's path breaking work is based on trends and relationships observed at the global level; in this study we make comparisons at the country level. Further, we provide a micro-foundational theory of movement based on individual decision making.

The paper is presented in four sections. We lay out the argument and hypotheses in the following section. Then we discuss the sample and operational indicators. We follow that with a discussion of our estimation and results. Finally, we discuss our plans for improving the analyses in the conclusion.

The Argument

We are interested in identifying the characteristics of countries that make them more likely to produce a greater proportion of refugees than IDPs. As such, we are interested in explaining a macro-level observable. Yet because these observables are the outcome of the decisions made by individual people, we desire an explanation rooted in individual decision making.[2] Because our problem—forced migration—is similar to the problem of voluntary migration, we begin with a simple model of voluntary migration. The idea that this simple model captures is that people decide to relocate on the basis of a comparison of the opportunities (e.g., future earning potential) available in their present location versus the opportunities available in an alternative location. If the opportunities elsewhere exceed the opportunities in the present location by a value greater than the costs of relocating, then the person will relocate. To fix ideas, we can write:

$$U_R = (B_R - C_R) - B_{\sim R}$$

where U represents utility, B represents benefit, C represents cost, R represents relocate, and ~ represents 'not.'[3] A person will relocate if $U_R > 0$. If $C_R > 0$, then B_R must exceed $B_{\sim R}$ by at least C_R before a person will relocate.

Because the decision to relocate as a forced—rather than a voluntary—migrant is different, we need to revise the model before it is useful for our study. People who abandon their homes due to a fear of persecution are making a choice under a considerably more constrained situation than people who choose to migrate

[2] See Little (1991, ch. 2) for a useful discussion of the value of micro-foundations in theory.
[3] We make this model as simple as possible. It could be complicated by making it an expected utility model within which we might think of the probability of receiving the benefits as p and 1-p. These probabilities could be thought of as driven by the persons' uncertainty, and it would be quite reasonable to assume that the a person would have greater uncertainty about benefits in the relocating place than in her current situation. We are implicitly assuming that p is .5, and thus drop it from the model. Further, we assume that the persons' discount rate on future benefits is constant across the two locations, and thus do not include a discount term.

in the absence of a fear of persecution, and the revised model must account for that difference. While the structure of the model remains much the same, the major difference is the information relevant to people we are modeling. In a voluntary migration model one would not include information about violence in the cost term, but in a forced migration model information about violence is the critical component of the cost term. Thus, we need to add an additional term—the cost of not relocating ($C_{\sim R}$)—to the model. We elaborate on this point below, but first discuss how we revise the simple voluntary migration model so that it can capture the choice between relocating within one's own country and relocating abroad.

To adapt this simple model to the choice between relocating in one's own country and relocating abroad we need to revise the notation as follows:

$$U_A = (B_A - C_A - TC_A) - (B_O - C_O) - (B_{\sim R} - C_{\sim R})$$

where TC represents transaction costs, the subscript $_A$ represents the decision to relocate in an asylum country, and the subscript $_O$ represents the decision to remain within ones' own country.[4] This revised decision model suggests that forced migrants are people who have compared the opportunities in their present location with those in alternative locations and decided to leave, but we have yet to identify the relevant information on which people will make their judgments. To use the model to produce hypotheses about the prevalence of IDPs relative to refugees in a given country we need to address several issues.

We begin by explaining how people assign values to the terms in our model. Our approach is conventional: people use information in their environment to assign values. More specifically, we assume that there is a single set of information about

[4] For convenience we assume that the person has already left her home and relocated within her country. Thus, we do not need to include a cost term for relocating in the origin country.

levels of violence, job opportunities, the level of democracy, etc. in a country: people have access to roughly the same information on these topics. Because preferences and resource allocations will vary across individuals, not everyone will respond the same way to that information set. However, large numbers of people—those whose preferences and resources are near the high point of the joint distributions—will respond to the information set in the same way. We wish to account for the migratory behavior of these people.

Let us be more specific about preferences and resources. We assume that the values of B_A, C_A, TC_A, B_O, and C_A, are normally distributed across individuals facing similar decisions and information.[5] That is, we assume that people have different preferences (i.e., assign different values to B) and different resource endowments (i.e., are affected differently by C and TC). This assumption is important only because without it our model would lead to the implication that all people (with similar personal circumstances) would respond in the same way to informational stimuli. That implication is false on its face as the different responses of Kosovars in 1999 indicate, and we invoke the assumption to avoid it.

Our assumption that people use information to assign values to the terms in our model is thus important for making the micro-macro link: by using information that is the same for all people in a country we can develop hypotheses about how people will respond to that information, and thus produce hypotheses about the macro-observable outcomes across different countries. Having invoked that assumption, we must specify the information that people will use to assign values to the B, C and TC terms.

[5] The normal distribution is merely convenient—many other distributions would serve our purposes equally well. The big issue is that they are not uniformly distributed.

Following Moore & Shellman (2002), which builds on Davenport et al (2002), we focus our attention on three groups of concepts: violence, socio-econo-political opportunity, and transaction costs. Because we need to distinguish these terms across asylum and origin countries, we present our hypotheses in two groups: local effects and neighborhood effects. First, however, we introduce the major concepts.

Violence

With respect to violence, we contend that two actors are likely to make people fear persecution (i.e., deprivation of life, liberty, or physical person): the state and dissidents. We argue that people will monitor the behavior of the state and the behavior of dissident groups to evaluate the value of $C_{\sim R}$. More specifically, the higher the level of violence, the higher the value of $C_{\sim R}$, and the more likely people are to abandon their homes as either refugees or IDPs. Thus, if the state engages in violence against its citizens, then we expect the value of $C_{\sim R}$ to rise across the population.[6] In addition, if dissident groups are engaged in violent attacks or a terror campaign, or groups in civil society engage in rioting, people will be more likely to re-evaluate $C_{\sim R}$. Like state violence, we expect dissident violence to increase the probability that people abandon their homes, either to become IDPs or refugees.

While we expect the level of violence by the state and dissident groups to increase the probability of observing non-zero levels of both refugees and IDPs, we expect the level of violence from both actors in potential countries of asylum to lower the probability that a country produces more refugees than IDPs and raise the

[6] Recall that we do not assume that these terms will rise uniformly across the population, but rather that the resulting distribution of, in this case, C_A terms will be approximately normal.

probability that a country produces more IDPs than refugees. Put in terms of the utility model, state or dissident violence in neighboring countries should lead people to increase the value of C_A but not the value of C_O. We thus hypothesize that local state and dissident violence will increase the probability that a country produces both IDPs and refugees, and that state and dissident violence in neighboring countries will decrease the probability of observing more refugees than IDPs, and increase the probability of observing more IDPs than refugees in that country.

In addition, we contend that when both groups engage in an armed contest (i.e., when a state of civil war exists) this will have an independent effect on the likelihood that people relocate. A civil war in the country of potential origin will lead people to raise their estimates of $C_{\sim R}$. Civil wars in neighboring countries, on the other hand, should lead people to increase the value they assign C_A, but not the value of C_O. This leads us to hypothesize that civil wars in neighboring countries will have a negative impact on refugee flows and a positive impact on IDP flows.

Finally, when there is an international war and battles take place on the territory of the country of potential origin, people will be more likely to feel threatened, and assign a higher value to $C_{\sim R}$. As with the other measures of violence, we expect that international war on the territory of neighboring countries will cause people to revise upward their beliefs about C_A, but not their beliefs about C_O.

Socio-econo-political Opportunity

Many migration models focus strictly on wages (or income) as a benefit. We certainly agree that income can be an important benefit, but we do not believe it is the only 'good' that people pursue. In addition, we believe that people value living in

a culture they find familiar and that they value the interpersonal relationships they have with family and friends. We further believe people will prefer political freedom and the rule of law to repression and the arbitrary exercise of authority. Thus, we anticipate that people will monitor information that enables them to make judgments about the expected benefits they could accrue on those dimensions in one location versus another.

With respect to income, we expect that people have access to information about the size of the economy, both in the potential country of origin and in neighboring countries. Size of the economy is a more general concept than wages, which is the focus of many migration models. Operational problems prevent us from focusing on wages,[7] but we also adopt size of the economy because (1) it is influenced by both the size of the country and population, and (2) we think that type of information is widely diffuse. The latter is important for obvious reasons, but we like the former because we are most interested in (perceived) economic opportunity. Large countries (in terms of territory) with large populations and large economies have—on average—more economic opportunities than smaller countries with smaller populations and smaller economies. Size of the economy is the concept that best captures these dimensions.

With respect to the utility model, we expect the size of the local economy to have a positive relationship with peoples' estimates of both B_O and $B_{\sim R}$, and a negative relationship with peoples' estimates of B_A. With respect to neighborhood effects, we hypothesize that the size of the economy in neighboring countries will

[7] Specifically, the absence of a cross-national time-series data base of comparative wages makes it difficult for us to focus on wages.

have the opposite impacts: negatively related to the values people assign to both B_O and $B_{\sim R}$, and positively related to the value people assign to B_A.

The next variable of interest is a familiar culture and family networks. It is well known that a key determinant of migration flows is a network of friends and family who have migrated (Faist 2000:40, *The Economist* 2002:6). Such a network not only creates a familiar culture in a new location, but also provides information which will lowers one's uncertainty about what to expect.[8] We thus expect that the larger the number of IDPs and refugees in/from a given country, the greater will be peoples' values for B_O and B_A. We thus hypothesize that a measure of the stock of past forced migrants will be positively associated with the probability of both producing IDPs and refugees. We do not have a hypothesis about whether the stock of past forced migrants produces more refugees than IDPs, or vice versa.

Turning to political freedom and the rule of law, we expect that the more democratic a country's political institutions, the higher will be peoples' assessment of the B_O term. The neighborhood level of democratic institutions should positively influence the value people assign to the B_A term. We hypothesize that the extent to which local institutions are democratic will be positively associated with the probability that a country produces more IDPs than refugees and negatively associated with the probability that a country produces more refugees than IDPs. With respect to the neighborhood's level of democratic institutions, we anticipate that it will be co-vary positively with the probability that a country produces more refugees than IDPs and negatively with the probability that a country produces more IDPs than refugees.

[8] This last point is more of a transaction cost argument, and one could contend that it lowers the TC_A term (Davenport, et al. 2002).

Transaction Costs

Finally, we assume that transaction costs, TC_A, are a positive function of distance and difficult terrain. To produce specific hypotheses we assume that relocating within one's own country requires less travel than relocating abroad.[9] This leads to the hypothesis that, ceteris paribus, people are more likely to relocate within their own country than abroad. With respect to difficult terrain we hypothesize that countries surrounded by mountains or with water barriers on their border will be more likely to produce less refugees than IDPs.

Research Design and Operationalization

To test our hypotheses we developed a sample using the country-year as our unit of observation. Our temporal domain is 1970-95. We use all countries for which we could obtain data on each of the variables that we used in the study.[10]

The Dependent Variable

To measure refugees we used a data set provided to us by Bela Hovy, the Director of the Statistics Division of the United Nations High Commission for Refugees. Susanne Schmeidl provided our IDP data—it is unpublished data from the Global

[9] We recognize that this assumption is violated for people who live closer to a city, town or village across the border than the nearest city, town or village in their own country.

[10] Known as complete case analysis, this sample selection suffers from bias to the extent that the cases that are missing data are not a random selected set of countries. To the extent that the selection of cases with complete data is systematic and in some way related to our study, then our inferences will be biased. One way to try to address this potential bias problem is to assign values to the missing cases. In an appendix we discuss the results we obtained using three different data sets composed of data that includes observations that are missing in the results reported in the paper.

Refugee Migration Project (Schmeidl and Jenkins, 1999). These are the best available sources of data on refugee and IDP stocks.[11] However, we are interested in flows, not stocks. That is, we want to know how many people relocated and became refugees or IDPs in a given year, not what the number of people with such status was in a given year. As Schmeidl (1998, 2000) explains, the annual refugee stock data are the most valid, reliable data available and there is virtually no large-N data on refugee or IDP flows. So one needs to work with the stock data to produce a flow measure.

When people speak of flow they generally mean the net flow, which is to say the number of people who have migrated minus those who have returned. We are interested in the former, but not the latter. The best, though imperfect, approach to measurement is to take the first difference of the stock (i.e., subtract the present year's value from the previous year's value), and then recode all negative values (which represent a net return flow) to zero.[12] We used this approach to create both a refugee flow and an IDP flow measure.

[11] The UNHCR data cover the period 1954-1999 and are the product of an extensive, multi-year effort by Hovy's team. The IDP data cover 1970-1997 and were put together by Susanne Schmeidl, working with Craig Jenkins. They are also the product of an extensive, multi-year effort (Schmeidl also worked with Hovy). See Schmeidl, (1998, 2000) and Crisp (nd, 2000) for detailed discussions about the strengths and weaknesses of these data. A key theme in Crisp's work is that different actors have different incentives to inflate or deflate the estimates of IDPs and refugees. Because we cannot say in a large data set like ours whether the average bias is positive or negative, we cannot judge whether the size of the effects we report are over-estimates, under-estimates, or about right.

[12] The weakness of this operationalization is that it will undercount flow in cases where a non-zero stock exists and a resettlement occurs. For instance, if 1,000 refugees are reported in year 1, but 800 are reported in year two, our measure will produce a value of 0 (the –200 will be recoded to 0). It may well be that 200 people returned, but it might also be the case that all 1,000 people were resettled in a third country and 800 additional refugees crossed the border. Thus, our measure likely undercounts flows and thus the size of the impact of our variables is smaller than they should be.

The dependent variable that ultimately interests us is whether a country that produced forced migrants produced more IDPs or more refugees. We therefore coded a nominal variable with three values:

0 indicates that the country produced neither IDPs nor refugees in that year

1 indicates that the country produced more IDPs than refugees in that year

2 indicates that the country produced more refugees than IDPs in that year

Measures of Violence

We use two different variables to measure state violence: genocide/politicide and violations of the rights of the integrity of the person. To measure genocide we use Barbara Harff's data (Harff and Gurr 1988, 1996). These data code events where the governing elites or their agents implicitly or explicitly promoted sustained policies that produced deaths in a substantial portion of the membership of a group of people. The variable, genocide, is coded 0 for country-years in which no such event occurred and 1 in country-years when such an event took place. Schmeidl (1997) and Davenport, et al. (2002) have shown that this variable has a positive impact on refugee stocks and net refugee flows.

We use the Political Terror Scale (PTS) data to measure human rights violations (Gibney & Dalton 1997).[13] The PTS is available for the years 1976-1996 and is a standards-based measure of the extent to which a government violate the physical right to integrity of the person. It uses a 5 point scale where higher values are associated with greater levels of violation. The data are collected via content

[13] Apodaca (1998) finds that government coercion—as measured via the Political Terror Scale—has a positive impact on refugee stock. Davenport et al. (2002) also use the political terror scale to measure government coercion, but fail to find support it impacting net refugee flows.

analysis of two sources: Amnesty International annual reports and the US Department of State's annual reports on human rights. Two different variables are created, one based on the Amnesty International coding and one based on the Department of State coding. We report the findings using the Department of State measure, but the results were substantially the same when we used the Amnesty International measure.

To measure dissident violence we use an event based measure and calculate the sum of riots and guerrilla attacks using the Cross-National Time-Series Data Archive (Banks nd). This is a revised version of a variable used in Davenport's (1995) study of government coercion. We dropped the two non-violent components of that measure—general strikes and anti-government protests—because these events are unlikely to lead people to conclude that their lives, liberty or person are threatened. The variable is a frequency count of these events in a given country-year.

Both government coercion and dissident violence may be episodic rather than sustained. In fact, this is generally the case. However, when both the government and the dissidents engage in more regularized military clashes, people are likely to have a heightened sense of insecurity, even controlling for government coercion and dissident violence. To measure this variable we use the Correlates of War Intrastate and Extra Systemic war lists (Sarkees 2000). It is a dichotomous measure scored one whenever more than 1,000 battle deaths occurred in a country-year and zero otherwise. We also use the Correlates of War list of interstate wars to help us code the presence of a war on the territory of the country of origin (Sarkees 2000). We examined the descriptions of the wars found in historical digests and an encyclopedia of war to determine whether at least one battle had taken place on the

territory of each country involved. When a battle took place in the territory, we assign the country-year observation a value of one. All other countries-years (including those not listed in the interstate war data) were assigned a value of zero.

Measures of Socio-Econo-Politico Environment

We identified three concepts that contribute to benefits: size of the economy; culture, family and friends; and freedom. We use gross national product (GNP) to measure the size of the economy. We obtained our GNP measure from both the World Bank's World Development Indicators data and the Cross-National Time-Series Data Archive (Banks nd). We began with the World Bank data and then filled in missing observations with the Banks data, where available.[14]

To measure the extent to which culture, family and friends are present, either in potential asylum countries or within one's own country, we used the lagged value of the stock of forced migrants (i.e., the sum of refugees and IDPs). The stock measures all of the people who had migrated—within their country or across a border—in the past, not simply those who had migrated the year before. This measures the size of the forced migrant population from each country, and thus the extent to which those who abandon their homes will find a community of people who share their culture, language, etc.

To measure freedom we use the Polity IV data. Specifically we subtract the autocracy score from the democracy score to create a measure of institutional democracy that ranges from –10 to 10. One of the challenges of the Polity data is

[14] The correlation between the World Bank's data and Banks' data is .918. In addition, when we use either of the series in a regression, the estimated coefficients are essentially the same. Given the strength of the correlations and the extent to which they produced similar estimates, we chose to replace missing data in the World Bank data with observations from the Banks' data.

that there are periods of transition in a country where the institutions are not sufficiently embedded for the project to produce a polity code. These cases show up in the data as having values of –66, –77, or –88. Rather than treat these as missing values we assigned them a score of 0 and coded a dichotomous measure of transition polities which we assign a value of 1 when the democracy score has a value of –66, –77, or –88, and a value of zero when democracy is scored in the range from -10 to 10.

Measure of Transaction Costs

We use the existence of mountains on a border to measure transaction costs. These are data coded by Shellman (2001). Mountain is a dichotomous measure coded one whenever at least 50 percent of a border has an elevation change of 1,000 feet or more from the surrounding area.

To develop an asylum neighborhood measure of both mountain and border we simply took the mean score across the borders for each country. The result is simply the proportion of neighboring countries that have mountains on their border. Thus, if a country had four borders and two with mountains, the value is four divided by two, which yields .5.

Ideally we would also have measures of both size of population in each region of the country, the location of violence, and distance to borders. This would allow us to include the distance people need to travel, a transaction cost that undoubtedly affects decisions. In a global analysis such as this, it would be very costly to develop such data, and it is probably not possible to develop data on the location of violence. As we discuss in the conclusion, in future studies we plan to conduct some time series case studies where we can develop such measures—the goal in this analysis is to determine whether the gross types of measures we use here can help us identify

characteristics of countries that are helpful for distinguishing refugee producers from IDP producers.

Neighborhood Scores

Because we expect the information about benefits and costs of seeking asylum to have a different effect on refugee and IDP flows, we need to measure the benefits and costs of seeking asylum in different countries. Doing so is not straightforward. The challenge involves drawing comparisons across countries which have different structures of information; that is, different numbers of neighbors. Since some countries have only one bordering neighbor, while others have two, three, four, etc., we cannot compare across a standard set of options.

To try to get a handle on the asylum versus origin decision, we create a summary measure to assess the 'asylum neighborhood.' We anticipate that the greater the benefits in the asylum neighborhood, the greater refugee flows will be, and the lower IDP flows will be. Similarly, we expect greater costs in the asylum neighborhood to be associated with lesser refugee flows and stronger IDP flows.

We create a neighborhood score for the following variables:

B_A: size of culture, GNP, democracy

C_A: genocide, rights violations, violent dissent, civil war, international war on territory,

We wish to develop a summary measure of the neighborhood on each of the above variables. More specifically, we want to know whether the country of origin is generally better (or worse) than the neighborhood on each of those variables. As such, we compare the score on each variable in the country of origin with the score

on the variable in each neighboring country, and then determine the proportion of countries that have a higher score.

To take a concrete example, consider a continuous measure, such as GNP. We begin with a directed dyadic data set where the observations are countries of potential origin and one of their neighbors. Thus, Afghanistan, which has six neighbors in 1995, is a country of potential origin six times, as follows:

Year	Origin	Asylum
1995	Afrghanistan	China
1995	Afrghanistan	Pakistan
1995	Afrghanistan	Iran
1995	Afrghanistan	Turkmenistan
1995	Afrghanistan	Uzbekistan
1995	Afrghanistan	Tajiskistan

Each of the above is an observation in the data, and when borders change, we code those changes in our data. Returning to the GNP example, we code a dummy variable that is assigned the value of 1 when GNP is higher in the country of potential asylum (e.g., Pakistan) than it is in the country of origin (e.g., Afghanistan). We then use the `collapse` command in Stata to calculate the proportion of neighboring countries that have a higher GNP for each potential country of origin. The `collapse` procedure calculates scores by group, which in our study is potential country of origin. After running the `collapse` command, we have a country year data base, and the variables produced by the `collapse` command range from 0 to 1, where a value of 0 indicates that the potential country of origin has a higher GNP than each of its neighbors and a value of 1 indicates that it has a

lower GNP than each of its neighbors. When the variable is dichotomous (e.g., our measure of civil war), then we do not take the intermediate step of creating a dummy variable: we simply use the `collapse` command on our dichotomous measure.

Estimation and Results

Because our dependent variable is a nominal, multichotomous measure, we chose the multinomial logit model to estimate the coefficients (Long 1997, ch 6), and used the `mlogit` command in Stata 7. These models produce n-1 sets of coefficients, where n is the number of nominal categories for the dependent variable. Since we have three nominal categories (no forced migrants; more IDPs than refugees; and more refugees than IDPs), we get two sets of estimated coefficients. The estimated coefficients indicate the log-odds of observing a given outcome in comparison with a baseline category. We use 'no forced migrants' as our baseline category and the `mlogit` procedure produced estimated coefficients for the 'more IDPs than refugees' value and the 'more refugees than IDPs' value. Because log-odds values are difficult to interpret, we also report the factor change values (aka relative risk ratios), [15] which have a straight forward interpretation.

In addition, we conduct some Wald tests to determine whether the estimated coefficients are sufficiently different across the 'more IDPs than refugees' and 'more refugees than IDPs' equations. Doing so helps us establish whether the independent variables have a different effect, and we expect that some of them will. Finally, we

[15] See Long (1997:168-70) for a discussion of factor change scores. In Stata 7 the `rrr` sub-command of the `mlogit` command produces these values.

report some predicted probabilities using different values for a handful of variables of interest. This helps us illustrate the substantive impact of those variables.

Table 1 reports the findings for two models. The first has a larger sample because it does not include our measure of human rights violations, which are only available beginning in 1976. The second model contains that variable, and has a smaller sample.

[Table 1 about here]

The results indicate that several of the variables have a statistically significant impact on both the probability that a country produces more IDPs than refugees, relative to a country that produces neither IDPs nor refugees, and the probability that a country produces more refugees than IDPs, relative to a country that produces no forced migrants. Further, many of those variables have a statistically different substantive effect on the probability of producing more IDPs than refugees and the probability of producing more refugees than IDPs. So at a very general level, the results indicate that while a single general process produces both IDP flows and refugee flows, we can identify factors that have a differential impact on the probability that a fear of persecution will produce more IDPs than refugees, and vice versa. Turning to specifics, we discuss our findings by focusing on clusters of variables.

Violence Indicators: Local & Neighborhood Effects

We begin with the violence variables, which measure the extent to which large numbers of people are likely to develop a reasonable fear of persecution. Consider the measures of state violence and dissident violence. We know from other studies that state-initiated violence and dissident-initiated violence increase forced migrant

flows (Schmeidl 1997, Davenport, et al. 2002, Moore & Shellman 2002). The results in Table 1 suggest that state-initiated violence and dissident-initiated violence have somewhat different effects on the probability that a country produces more IDPs than refugees, and vice versa. First, note that our measure of genocide produces a statistically significant estimate for the 'refugees greater than IDPs' equations, but not for the 'IDPs greater than refugees' equations. Further, the estimated coefficient is substantively larger in the 'refugees greater than IDPs' equation in both models. A Wald test indicates that—for both model 1 and model 2—we can reject the null hypothesis that the coefficients have the same value in both equations. These findings suggest that genocidal events increase the probability that a country will produce more refugees than IDPs, but that they do not affect the probability that a country will produce more IDPs than refugees. More specifically, the odds that a country produces more refugees than IDPs relative to one that produces no forced migrants is between 1.96 and 2.43 times greater for countries that experience a genocidal event than for countries that do not.

Turning to the political terror scale,[16] our measure of human rights violations, we see that the estimated coefficient is statistically significant in both equations. Human rights violations have a larger impact on the probability that a country produces more IDPs than refugees, but the difference is rather small (a Wald test shows that the difference is statistically significant).[17] The estimated coefficients for our measure of dissident violence tell a similar story: three of the four

[16] Because it yields a slightly larger sample, we report the results obtained when we used the US State Department measure. The results obtained using the Amnesty International measure are effectively the same except where mentioned in a footnote.
[17] Further, the difference between the two coefficients for the political terror scale variable is not statistically different from one another when we use the Amnesty International measure (though both coefficients are statistically significant).

are statistically significant, and Wald tests indicate that the estimated coefficients are not meaningfully different in value across the 'IDPs greater than refugees' and the 'refugees greater than IDPs' equations.

Let us now consider the neighborhood measures of state and dissident violence. The first thing to note is that none of the neighborhood measures produce a statistically significant coefficient in the 'IDPs greater than refugees' equations: the extent to which a country's neighbors exhibit greater state or dissident violence than the potential country of origin has no impact on the probability that the country produces greater IDPs than refugees. However, the extent to which the neighborhood is violent does have an impact on the probability that a country produces more refugees than IDPs: our measures of both genocide and dissident violence produce negatively signed parameter estimates, and three of the four are statistically significant. Given the non-linearity of the parameters, the negative signs may be misleading, so note that the relative risk rates are less than 1. For example, the odds that a country produces more refugees than IDPs is between 0.31 and 0.17 times lower relative to a country that produces no forced migrants if all of its neighbors are experiencing a genocide than if none of its neighbors are experiencing a genocide. Wald tests indicate that the estimated coefficients for the proportion of genocide are different across the two equations for model 1, but not model 2. Further, we cannot reject the null hypothesis that the coefficients for the proportion of dissident violence in neighboring countries are the same. The proportion of neighboring countries with a higher level of human rights violations does not have a statistically significant impact in either equation.

We now consider our measure of the interaction of states and dissidents above a threshold of violence—civil war—as well as our measure of foreign troops

fighting on a country's soil. Civil war produces statistically significant estimates for both equations in both models. Note further that the substantive impact of civil war is markedly different (Wald tests demonstrate that the difference is statistically significant). Specifically, a country that experiences a civil war is between 11.5 and 25.9 times more likely to produce more IDPs than refugees relative to a country that produces no forced migrants than a country that does not experience civil war. By comparison, a country undergoing civil war is between 1.82 and 2.85 times more likely to produce more refugees than IDPs relative to a country that produces no forced migrants than a country that does not experience civil war. Thus, while civil war increases the probability of observing forced migrants in general, it has a considerably larger impact on the probability that a country produces more IDPs than refugees than the reverse.

An international war fought on a country's territory has a similar effect: it increases the probability that a country produces more IDPs than refugees. This variable does not produce a statistically significant parameter estimate for the 'refugees greater than IDPs' equation in either model 1 or model 2.[18] A country that experiences war battles is between 3.7 and 4.35 times more likely to observe more IDPs than refugees relative to no forced migrants than a country that does not have battles fought on its territory. The results suggest that battles fought as part of an international war have a greater impact on IDP production than refugee production.

What impact do civil war and international war in the neighborhood have? The results in table 1 suggest that the proportion of neighboring countries with international war battles being fought does not affect the relative production of IDPs and refugees relative to no forced migrants. The proportion of civil war among

[18] The Wald tests indicate that the coefficients are not different across the equations.

neighbors, on the other hand, does increase the probability that a country will produce more IDPs than refugees, and this finding is consistent with our expectations. More specifically, a country whose neighbors are all experiencing civil war is between 5.33 and 21.6 times more likely to produce a greater number of IDPs than refugees relative to a country that produces no forced migrants than one which has no neighboring countries engaged in civil war. With respect to the second equation, with 'refugees greater than IDPs' we expected the proportion of neighboring countries with civil wars to have a negative impact. We cannot reject the null hypothesis for a one-tailed test in either model 1 or model 2 for that equation. It is surprising that the coefficient for the proportion of civil war variable is both positive and greater than twice its standard error in model 1 (it is positive, but considerably smaller than its standard error in model 2).

One ad hoc explanation for this finding is suggested in Weiner's (1996) article on bad neighborhoods. Weiner suggests that countries with civil wars cluster together across the globe. This suggests that it is possible that countries with civil wars tend to be next to other countries with civil wars, and implies that our finding may be an artifact of unmodeled endogeneity (specifically, diffusion of civil war). The design of this study cannot rule that out, but the correlation between civil war and the proportion of neighbors with civil wars, .05, casts doubt on this account. In any case, there are likely many other ad hoc accounts for the finding. All we can infer, however, is that our hypothesis that countries in violent neighborhoods will have a greater probability of producing more IDPs than refugees is not supported by these data.

Socio-Econo-Politico Indicators: Local & Neighborhood Effects

The first indicator we consider is the lagged value of the stock of forced migrants from a given country, which we use as our measure of the extent to which a familiar culture exists in potential migration locations, both within and outside the given country. This variable has a statistically significant and positively signed coefficient in the 'IDPs greater than refugees' equation in both models and the 'refugees greater than IDPs' equation in model 1. The relative risk rate score of 1.00 is not very helpful, but this is not surprising given that one would not expect a one unit increase—which is the basis of the `rrr` score—to have much of an impact, and a value of 1.00 indicates no change in the probability. Given that the lag of the stock of forced migrants has a range in the millions, it is more useful to calculate the change in the predicted probability over a range of values (this is true of all independent variables with a large range). To consider the largest possible effect, holding all other variables at their mean, a change from the minimum to the maximum value of the lag of forced migrants increases the probability of observing more IDPs than refugees by .07 in model 1 and .02 in model 2, and increases the probability of observing more refugees than IDPs by .39 in model 1 and .10 in model 2.[19] These results demonstrate that the variable has the positive impact that we expected: the larger the group of forced migrants living elsewhere as IDPs or refugees, the greater the probability of observing forced migrant flows, regardless of whether there are more IDPs than refugees, or more refugees than IDPs.

Next we consider our measure of economic opportunity, GNP. The results in table 1 show that our hypothesis is supported with only one of the four estimated

[19] We must note that this last coefficient is not statistically significant.

coefficients: GNP is negatively signed and statistically significant in the 'refugees greater than IDPs' equation in model 1. We hypothesized that the larger the GNP at home, the lower the probability that a country would produce more refugees than IDPs, and this estimated coefficient appears consistent with the hypothesis. To determine whether it is we must look at predicted probabilities as a one dollar increase in GNP will not have an impact on the probability of producing more refugees than IDPs, as the relative risk rate of 1.00 indicates. A change from the minimum to the maximum GNP, holding all other variables at their means, decreases the probability of observing more refugees than IDPs by .10. Interestingly, however, the estimated coefficient for GNP is also negative in the 'IDP greater than refugee' equations for both model 1 and model 2. Further, the estimated coefficient is more than twice its standard error. As before, we can only infer that our hypothesis of a positive relationship between GNP and the probability of observing more IDPs than refugees is inconsistent with this data. One plausible ad hoc explanation is that poor countries are more likely to produce forced migrants (i.e., both IDPs and refugees) than well to do countries, and that effect simply washes out any impact that economic opportunity might have on the relative levels of IDPs versus refugees in any given country.[20] Other ad hoc explanations might provide similarly plausible accounts for the findings.

Let us now consider the neighborhood effect of economic opportunity. The neighborhood GNP variable measures the proportion of bordering countries that have a higher GNP than the potential country of origin. The results in table 1 indicate that economic opportunity in the neighborhood has a statistically

[20] Unfortunately, the Wald test—which would provide some evidence to support or challenge this explanation—would not compute, presumably because the size of the coefficients were so small.

significant impact on the probability that a country produce more refugees than IDPs, but not the probability that a country produces more IDPs than refugees. The relative risk rates indicate that a country that is surrounded by neighbors with larger economies is 1.39 to 1.75 times more likely to produce more refugees than IDPs relative to a country the produces no forced migrants than a country that is surrounded by neighbors with smaller economies. This finding is consistent with our hypothesis that the stronger the economies in the neighborhood, the greater the likelihood that refugee flows will be larger than IDP flows. Yet, it is interesting to note that the size of the impact is fairly small: a country that has a smaller economy than all of its neighbors is 1.39 to 1.75 times more likely than one that has a larger economy than all of its neighbors to produce more refugees than IDPs. Given that most countries have several neighbors, the change in the independent variable is rather large, and the corresponding change in the odds, while not inconsequential, is not especially large. Thus, large neighboring economies do draw refugees, but the effect is modest.

The findings for political freedom—as measured by the Polity data's democracy-autocracy scale—are consistent with our hypotheses. All four estimated coefficients are statistically significant and in the expected direction. The relative risk rates indicate that the probability of a country generating more IDPs than refugees relative to producing no forced migrants is 1.04 to 1.06 times higher for every one unit increase in democracy. Further, the probability that a country produces more refugees than IDPs relative to a country that produces no forced migrants is between 0.98 and 0.94 times lower for every unit increase in democracy.

The change in predicted probabilities over the range of the democracy-autocracy scale make the case even more clearly: holding all other variables at their

mean and moving democracy from its lowest to its highest value leads to between a .01 and .02 increase in the probability of observing more IDPs than refugees and between a -.05 and a -.10 change in the probability of observing more refugees than IDPs. These changes appear small, but they are not when one puts them in the context of the baseline expected probabilities, which are .01 for 'IDPs greater than refugees' and .09 'refugees greater than IDPs.' Thus, a shift from the low to high score on democracy leads to between a 100% and 200% increase in the probability of observing more IDPs than refugees relative to producing no forced migrants. The same shift in democracy leads to between a 50% and 100% reduction in the probability of producing more refugees than IDPs relative to producing no forced migrants. While these effects are relatively small in comparison with some of the local violence indicators (e.g., civil war), they are still substantial.

What of the neighborhood effect of political freedom? Table 1 indicates that that the data are consistent with one of our hypotheses: the greater the proportion of neighbors with more democratic political structures, the greater the likelihood that a country will produce more refugees than IDPs. The coefficients are statistically significant for both of the 'refugees greater than IDPs' equations. In addition, the relative risk rates indicate that the impact is fairly substantial: a country that has a lower democracy score than all of its neighbors is between 3.15 and 3.61 times more likely to produce more refugees than IDPs relative to a country that produces no forced migrants than a country that has a higher democracy score than all of its neighbors. The estimated coefficient for the 'IDPs greater than refugees' equations are not consistent with our expectation that the more democratic the neighborhood, the lower the probability that a country will produce more IDPs than refugees. The

coefficient for model 1 is not only positive, but is greater than twice its standard error. We have not been able to produce an ad hoc explanation for this finding.

Transaction Costs: Neighborhood Effects

Our measure of transaction costs is the proportion of neighboring countries that have a mountain range on the border. Table 1 indicates that the data are inconsistent with our hypothesis: none of the estimated coefficients are statistically significant and properly signed. The estimates for the 'IDPs greater than refugees' equations are positively signed, but cannot be distinguished from zero with sufficient confidence. The estimates for the 'refugees greater than IDPs' equations are positively signed, contrary to our expectation that mountains increase transportation costs and would thus lower the probability that a country would produce more refugees than IDPs. Further, the estimates are more than twice their standard errors. A plausible ad hoc explanation for this finding can be found in Fearon & Laitin's (2001) study of insurgency, which emphasizes rough terrain as one of the most important determinants of the probability of an insurgency in a given country. Building on the insurgency literature that emphasizes hit-and-run tactics, Fearon & Laitin find that countries that do not have rough, inaccessible terrain—such as mountains—are very unlikely to develop an armed insurgency. Thus, our measure may be capturing hospitable conditions for insurgency as well as transaction costs. However, the correlation between civil war and mountains is -.07, which casts doubt on that explanation. In any case, the most we can infer is that our hypothesis about the impact of mountains on the probability that a country produces more refugees than IDPs is inconsistent with the data.

Discussion: Comparisons Across Variables

To this point we have discussed the findings in groups, ignoring comparisons across groups. These types of comparisons are important, not only with respect to making a small contribution to contingency planning, but also with respect to making a contribution to policy debates. Thus, we begin this section by making several comparisons across the variables. Having done so, we turn our attention briefly to contingency planning and the debate about economic motives for seeking asylum.

Table 2 is a two by two listing of the local and neighborhood variables that have a different impact on the probability that a country produces more IDPs than refugees relative to no forced migrants, or vice versa. Two variables that produced a statistically significant estimate—human rights violations and dissident violence—are not listed because they had essentially the same (positive) impact on both 'IDPs greater than refugees' and 'refugees greater than IDPs.' Table 2 lists the variables that had a different impact across the two equations. With respect to local variables, democracy is positively associated with a greater probability of observing more IDPs than refugees and negatively associated with the probability of observing more refugees than IDPs. Civil war and war on territory increase the probability of observing more IDPs than refugees, and genocide increases the probability of observing more refugees than IDPs.

[Table 2 about here]

Interestingly, there are no neighborhood variables that had an impact on the probability that a country produced more IDPs than refugees: in these data the neighborhood only influenced the probability of observing more refugees than IDPs. More specifically, the proportion of neighboring countries that are more democratic and the proportion that have a larger economy are each positively associated with a

greater probability of producing more refugees than IDPs. With respect to violence, the greater the proportion of neighboring countries that are experiencing a genocide event, the lower is the probability that the country will produce more refugees than IDPs. In the remainder of this section we discuss the implications of these findings with respect to two policy issues: contingency planning and the debate about 'economic refugees.'

Contingency Planning

Contingency planning involves developing scenarios for possible future events and making plans for responding to those events. With respect to complex humanitarian emergencies that could lead to IDP and/or refugee flows, contingency planning focuses on anticipating such flows, verifying the existence of adequate supplies and personnel, and perhaps establishing communication with the governments with which an agency will need to negotiate to provide humanitarian assistance in the event that people begin to relocate.

The first thing to consider is a simple bureaucratic politics, or t-1, model where we use the value from last year as a provisional estimate for the value this year. While the lagged value of the stock of forced migrants used in our analyses is not directly relevant because it does not distinguish IDPs from refugees, in separate analyses (not reported here) we found that the lagged value of the stock of IDPs has a statistically significant, positive, and substantively large impact on IDP flow. Similarly, the lagged value of the stock of refugees has a statistically significant, positive, and substantively large impact on refugee flow. Further, lagged stock of IDP did not produce a statistically significant parameter estimate for refugee flow, nor did lagged stock of refugees produce a statistically significant parameter

estimate on IDP flow. All of this suggests that the 'look at what happened last year' rule of thumb is useful.

But can we say more than that? The results reported in table 2 suggest that we can. Let us begin with the neighborhood. Neighborhoods with larger economies and neighborhoods with more democratic polities are associated with a greater likelihood of producing more refugees than IDPs. Thus, contingency planners for organizations that provide services for refugees may be well advised to consider the extent to which neighboring countries are more democratic and have larger economies than a country where violence is expected. Further, given that the relative risk rates indicate that the proportion of neighboring countries that are more democratic has between a 1.8 and a 3.6 times greater impact[21] on the probability of producing more refugees than IDPs than the proportion of neighboring countries with larger economies, contingency planners will want to weigh the proportion of more democratic countries more heavily than the proportion of countries with larger economies.

The other neighborhood effect of interest to contingency planners for agencies that service refugees is the proportion of neighboring countries that are experiencing genocide. The findings in table 1 indicate that the larger the proportion of neighboring countries experiencing genocide, the lower the probability that the country will produce more refugees than IDPs. The relative risk rates indicate that the impact is substantial. However, countries are extremely unlikely to be surrounded by neighbors experiencing genocide, so the relative risk rate is unrealistic. We can get a better sense of the impact by considering the marginal

[21] We arrived at these values by subtracting the relative risk rate (rrr) for NbhdGNP from the rrr for NbhdDemoc in each 'refugees greater than IDPs' equation.

effect, which is the partial derivative of the predicted probability with respect to an independent variable of interest, holding the other variables constant (Long 1997:165). The marginal effect for genocide at its mean level (which is a more realistic value to use than the maximum) is -.15, which demonstrates that the effect is substantial. Thus, the presence of a genocide in neighboring countries should lead those responsible for contingency plans to expect a lower likelihood that a country will produce more refugees than IDPs.

Turning toward local variables, we note that genocide in a country has a positive—and substantial—impact on the probability of observing more refugees than IDPs. The relative risk ratios of 1.96 (model 2) and 2.43 (model 1) indicate that a country faced with genocide is substantially more likely to produce more refugees than IDPs relative to producing no forced migrants. That genocide does not have a similar effect on producing more IDPs than refugees relative to producing no forced migrants suggests that those who create contingency plans for servicing new refugee populations will want to monitor genocide events.

The extent of democratic institutions is one of the more interesting variables in our study as it is positively associated with the probability of producing more IDPs than refugees and negatively associated with the probability of producing more refugees than IDPs. The relative risk ratios indicate that the effects are substantial: a one unit increase is associated with only a 4-6% increase in the probability of more IDPs than refugees and a 2-6% decrease in the probability of more refugees than IDPs, but the democracy scale has twenty one points, and countries can move a substantial distance on the scale in a brief period of time during transitions to or from democracy. So our results suggest that the extent of democratic institutions of

the local country have a substantial impact and warrant the scrutiny of contingency planners.

While the extent of democratic institutions is important, the measures that have the largest substantive effects are the two war measures, each of which influences the probability of producing more IDPs than refugees relative to producing more forced migrants, but not more refugees than IDPs. Thus, these variables warrant special attention from those contingency planners who work for agencies that serve IDPs. Further, civil war—which is measured using a battle deaths threshold of 1,000—has a considerably larger impact on the probability of 'IDPs greater than refugees' than does international war on local territory. Civil war also has a positive impact on the probability of producing more refugees than IDPs, but the relative risk ratios indicate that the impact of civil war, while substantial, is considerably smaller than it is for 'IDPs greater than refugees.' Thus, while both have a positive impact, the size of the impact is rather different. Civil wars are associated with both an increased probability of more IDPS than refugees relative to no forced migrants and more refugees than IDPs relative to no forced migrants. But the effect on the former is considerably stronger than the effect on the latter, implying that while contingency plans for both IDP and refugee flows are warranted, more often than not, civil wars will be associated with larger IDP flows than refugee flows.

To summarize, table 2 identifies variables that should prove useful if monitored by contingency planners responsible for servicing new IDP and refugee populations. The local violence measures have the biggest impact, and different ones affect the two outcomes in different ways. Local econo-politico variables also have an effect on both, but again in different ways. Finally, neighborhood violence,

economic size, and democratic structure each influence the probability of refugee flows being larger than IDP flows, but not vice versa.

Economic Refugees?

In a recent survey on migration *The Economist* offered the following opinion:

> Many asylum-seekers are fleeing poverty and disorder rather than persecution. It is usually impossible to establish whether they have really suffered the traumas they claim (November 2nd 2002, p. 7).

This is a widely held view and informs public debate about asylum policies across Europe today and was a focus of public discussion on the topic in the United States during the Reagan administration. While our analysis does not offer definitive evidence one way or the other with respect to that claim, it is of some interest.

Certainly there is no debate about the claim that some who seek asylum abuse the international refugee regime with trumped-up claims of a fear of persecution. The more interesting—and debatable—issue concerns the size of the proportion of asylum-seekers who harbor a legitimate fear of persecution compared with the proportion who do not. Is the proportion of 'economic refugees' large, as *The Economist* intimates, or is it relatively small? Surveys of individuals are not feasible, both because of prohibitive costs as well as the fact that people have an incentive to lie. And while it is notoriously difficult to infer individual motives from aggregate behavior, we note that the findings reported here are not very consistent with the assertion that a large proportion of asylum-seekers are 'economic refugees.'

We do indeed find, as *The Economist* and others would anticipate, neighborhood measures of both democracy and size of the economy are positively associated with the probability of producing more refugees than IDPs. However,

these effects are relatively small, and they are very small in comparison to measures of local violence. To wit, using the values from model 1 in Table 1 and holding all other variables at their means, a change from being the poorest country in the neighborhood to the richest country in the neighborhood produces a .025 increase in the probability of observing more refugees than IDPs. In comparison, shifting from a country without a genocide to a country with a genocide increases that probability by .10, and moving from an absence of civil war to the presence of civil war increases it by .09. As noted above, it is important to keep in mind that forced migration flows are low probability events, and these increases are in comparison to a baseline probability of .09.[22] Stated differently, either a genocide event or a civil war doubles the probability of observing more refugees than IDPs, and those are the events of concern to people who decry the abuse of the international asylum regime. Being surrounded by countries with larger economies, however, increases the probability of producing more refugees than IDPs by 28% (from .09 to .115).

At first glance, this is an impressive figure. Yet, one must consider two additional pieces of information that put the figure in context. First, a 28% increase occurs when we shift from being the wealthiest country in the neighborhood to the poorest country in the neighborhood, which is to say it is the extreme change. If we consider a one standard deviation change in the proportion of neighboring countries with larger economies, the change in predicted probability drops from .025 to .008, which is a modest 9% increase in probability. Second, and more importantly, when we compare the impact of size of the economies in the neighborhood to the impact of

[22] This figure is calculated with all variables held at their mean, model 1 in table 1. A very similar value is produced if we used model 2 in table 1.

local civil war or genocide, we see a sharp difference: having either a civil war or a genocide increases the probability by 100% and 111%, respectively.[23]

We have focused on size of the economy as that variable most closely represents the 'economic refugee' arguments. But we briefly observe that one reaches similar conclusions by examining the neighborhood effects of democracy. In fact, and this is perhaps surprising to advocates of the 'economic refugee' position, democracy is more of a draw than a large economy: a one standard deviation increase in the proportion of countries with more democratic institutions increased the probability of producing more refugees than IDPs by .03, which is a 33% increase (compared with a 9% increase for size of the economy).

Put simply, these figures are considerably more consistent with the claim that some asylum-seekers are opportunists seeking better economic opportunity, but the substantial majority of asylum-seekers are fleeing violence. We do not wish to overstate our case. First, our results provide only very rough estimates and do not help us settle 'how many is many.' Second, these are aggregate level data that do not permit direct inferences to the motives of individuals. Yet, to the extent that they can shed light on the issue, the findings cast doubt on claims that most asylum seekers are merely 'economic refugees' by showing that size of the economy has a considerably smaller impact on the probability that a country produces more refugees than IDPs than the extent to which a country is surrounded by more democratic countries. Further, the impact of democracy is considerably smaller than the impact of local violence. While we cannot answer the question 'how many is

[23] Again, these figures are based on the results from model 1, but the results using model 2 lead to the same conclusions.

many,' these results are considerably more consistent with the claim that most are fleeing violence than the claim that most are seeking opportunity.

Conclusion

We began this study wondering what could explain why several hundred thousand Kosovars abandoned their homes and relocated inside Kosovo and virtually none sought asylum abroad in 1998, but in 1999 near one million Kosovars sought asylum abroad while several hundred thousand sought safety within Kosovo. Our study does not provide a direct answer with respect to the Kosovo case. Instead, it sheds light on general patterns using global data from 1970-1995. We revise a simple migration model to make it relevant to the study of forced migration. It posits microfoundations of aggregate human behavior and focuses our attention on three sets of information: violence, socio-econo-political opportunity, and transaction costs. The latter two are relevant for models of voluntary migration; violence is a necessary informational component of models of forced migration.

Our findings identify several general patterns. First, both local violence and neighborhood violence have an effect on forced migration, but have different effects rather than a uniform effect. For example, whereas war on the territory of the country of potential origin substantially increases the likelihood that a country produces more IDPs than refugees, genocide has a positive influence on the likelihood of a country producing more refugees than IDPs. Civil war has a positive influence on both, relative to a country that produces neither, but a substantially stronger impact on the likelihood of producing more IDPs than refugees. Second, the extents to which government institutions are democratic positively influences IDP production and negatively influences refugee production. Third, neighborhood

effects do not have a systematic impact on the likelihood of producing more IDPs than refugees relative to countries that produce no forced migrants. Yet, fourth, econo-political neighborhood variables (size of the economy and democratic institutions) increase the likelihood of producing more refugees than IDPs, and a violence variable (genocide) decreases that likelihood. As we argued above, these findings have implications both for contingency planning and the debate about 'economic refugees.' That said, we turn our attention to the question: what additional work on this topic remains to be done?

The biggest weakness of global analyses such as this is the lack of more specific data about population locations (and, thus, distance from borders), their size, and composition (e.g., ethnic, linguistic, or racial groups). We believe that our lagged stock variables go a long way toward helping us find patterns of ethnic ties across borders, but they do not help us distinguish the sizes of targeted populations—our study effectively assumes that the probability of persecution or victimization is uniform across the population of each country, and this surely is not so. Further, we do not have any information about the location of violence, and it presumably makes a difference since people presumably (1) walk away from, not toward violence, and (2) the size and location of the population that is proximate to violence vis-a-vis borders will presumably affect the size of refugee flows. Finally, while we have some information about the geography of borders, we do not know anything about the terrain within countries, and this likely will affect decisions as well.

Another issue comes from consideration of Schmeidl's (1998) analyses of global trends. Her study reports change that has taken place over time in global patterns of both refugee and IDP stocks. Those trends might also be present in a

substantial portion of our countries, and thus we would do well to relax the assumption that we have constant parameters over time and explore the possibility that they may vary.

Finally, in future analyses we plan to design and carry out both case control design studies (where one holds important variables constant through sample selection) and time series case studies. The first approach is valuable in that we may be able to control for some variables we cannot measure to see whether the patterns in the global data change when we control for those variables. The second approach is critical because the costs of data collection change radically which will enable us to both collect data on concepts that we simply cannot measure in the global sample as well develop more reliable and more valid measures of concepts that we have measured. We can also reduce the level of temporal aggregation, and doing so will make it possible to explore causal relations more directly.

Given these shortcomings one might ask why we would begin with a global analysis. The reason is simple. We are interested in establishing what—if any—general patterns can be established. Doing so is useful for several reasons. First, the data we are using are easily collected and if they prove useful for out of sample forecasting (something we have not yet investigated), then we may have a rather crude, yet simple tool that will have some utility for contingency planning. Second, it gives us a baseline of knowledge against which we can assess the findings from the case control design studies and the time series case studies. Without the global analyses one would be hard pressed to assess the extent to which ones' findings were the result of selecting unusual cases. Third, these sorts of patterns can help illuminate new questions and, ultimately, answers as well as shortcomings not revealed by less systematic empirical investigations. Yet more fine-grained analyses

with both variables that cannot (yet) be measured in a global analysis as well as more reliable and valid measures of variables we can measure in a global analysis are needed. Once we complete our global analyses we will design and execute such analyses and determine the extent to which they further illuminate our understanding of forced migration flows.

References

Apodaca, Clair. 1998. "Human Rights Abuses: Precursor to Refugee Flight?" *Journal of Refugee Studies*, 11(1): 80-93.

Banks, Arthur S. nd. *Cross-National Time-Series Data Archive*, http://www.databanks.sitehosting.net/www/main.htm , accessed February, 2002.

CNN. 1999. "The Valley," CNN Perspectives.

Crisp, Jeff. 1999. "Who Has Counted the Refugees? UNHCR and the Politics of Numbers," *New Issues in Refugee Research*, Working Paper #12, Geneva: UNHCR, http://www.unhcr.ch/refworld/pub/wpapers/wpno12.htm, accessed July, 1999.

Crisp, Jeff. 2000. "Who Has Counted the Refugees? UNHCR and the Politics of Numbers," in S.C. Lubkemann, L. Minear and T.G. Weis (eds.) *Humanitarian Action: Social Science Connections*, Occasional Paper #37, Thomas J. Watson, Jr. Institute for International Studies, pp. 33-62.

Davenport, Christian. 1995. "Multi-Dimensional Threat Perception and State Repression: An Inquiry into Why States Apply Negative Sanctions," *American Journal of Political Science*, 39:683-713.

Davenport, Christian, Will H. Moore & Steven C. Poe. 2002. "Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964-1989," *International Interactions*, forthcoming; available at http://garnet.acns.fsu.edu/~whmoore/researchpapers.html, accessed January 2002.

Faist, Thomas. 2000. *The Volume and Dynamics of International Migration and Transnational Social Spaces*, Oxford: Clarendon Press.

Fearon, James D. & David D. Laitin. 2001. "Ethnicity, Insurgency, and Civil War," presented at the annual meeting of the American Political Science Association, San Francisco.

Gibney, Mark, Clair Apodaca & James A. McCann. 1996. "Refugee Flows: The Internally Displaced and Political Violence (1980-1993): An Exploratory Analysis," in A.P. Schmid (ed.) *Whither Refugee? The Refugee Crisis: Problems and Solutions*, Leiden: PIOOM.

Gibney, Mark & Matthew Dalton. 1997. "The Political Terror Scale," in D. Cingranelli (ed.) *Human Rights and Developing Countries*, Westport: Greenwood.

Harff, Barbara and Ted Robert Gurr. 1988. "Toward Empirical Theory of Genocides and Politicides: Identification and Measurement of Cases Since 1945," *International Studies Quarterly*, September, 32(3): 359-71.

Harff, Barbara and Ted Robert Gurr. 1996. "Victims of the State: Genocides, Politicides, and Group Repression from 1945 to 1995," in Albert J. Jongman (ed.) *Contemporary Genocides: Causes, Cases, Consequences*, Leiden: PIOOM, pp. 33-58.

Little, Daniel. 1991. *Varieties of Social Explanation: An Introduction to the Philosophy of Social Scien*ce, Boulder: Westview Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks: Sage.

Moore, Will H. and Stephen M. Shellman. 2002. "Fear of Persecution: A Global Study of Forced Migration, 1952-1995," working paper, available online: http://garnet.acns.fsu.edu/~whmoore/researchpapers.html, accessed October 2002.

Reliefweb. 2002. Online: http://www.reliefweb.int/mapc/eur_bal/reg/natorefidp.html, accessed July, 2000.

Sarkees, Meredith Reid. 2000. "The Correlates of War Data on War: An Update to 1997," *Conflict Management and Peace Science*, 18/1: 123-144.

Schmeidl, Susanne. 1997. "Exploring the Causes of Forced Migration: A Pooled Time-Series Analysis, 1971-1990," *Social Science Quarterly*, June, 78(2): 284-308.

Schmeidl, Susanne. 1998. "Comparative Trends in Forced Displacement: IDPs and Refugees, 1964-96," in J. Hampton (ed.) *Internally Displaced People: A Global Survey*, London: Earthscan Publications, pp. 24-33.

Schmeidl, Susanne. 2000. "The Quest for Accuracy in the Estimation of Forced Migration," in S.C. Lubkemann, L. Minear and T.G. Weis (eds.) *Humanitarian Action: Social Science Connections*, Occasional Paper #37, Thomas J. Watson, Jr. Institute for International Studies, pp. 127-57.

Schmeidl, Susanne and J. Craig Jenkins. 1999. *Global Refugee Migration Project*, http://www.soc.sbs.ohio-state.edu/jcj/research.htm, accessed March, 2002.

Shellman, Stephen M. 2001. *Global Directed Dyad Geographic Borders Time Series Database, 1964-1999*. http://garnet.acns.fsu.edu/~sms6202/data.html.

The Economist. 2002. "The Longest Journey: A Survey of Migration," November 2nd.

USCR. 1999. *World Refugee Survey, 1999*, Washington, DC: U.S. Committee for Refugees.

Weiner, Myron. 1992/93. "Security, Stability, and International Migration," *International Security*, Winter, 17(3): 91-126.

Weiner, Myron. 1995. *The Global Migration Crisis: Challenges to States and Human Rights*, New York: Harper Collins.

Weiner, Myron. 1996. “Bad Neighbors, Bad Neighborhoods: An Inquiry into the Causes of Refugee Flows,” *International Security*, Summer, 21(1):5-42.

Zolberg, Aristide R., Astri Suhrke & Sergio Aguayo. 1989. *Escape from Violence: Conflict and the Refugee Crisis in the Developing World*, New York: Oxford University Press.

Appendix

In this appendix we report several sensitivity analyses, the purpose of which is to evaluate the extent to which the findings reported in the text are robust. A concern with sample selection bias induced by missing data led us to use three techniques to generate data for missing observations: interpolation; interpolation and extrapolation; and imputation. We then re-estimated the equations reported above using the three augmented data sets to determine whether the findings were stable across the augmented data. In addition, we adjusted the standard errors using three corrections: robust standard errors, corrected standard errors clustered on country, and corrected standard errors clustered on a spatial variable (following _____, Ward & Gelditsch 2002).

This appendix remains to be written…

Table 1: Multinomial Logit Estimates,
IDP v Refugee Producers, 1970-1995

	Model 1				Model 2			
	IDPs > Refugees		Refugees > IDPs		IDPs > Refugees		Refugees > IDPs	
Variable	Coef (SE)	RRR	Coef (SE)	RRR	Coef (SE)	RRR	Coef (SE)	RRR
Genocide	0.23 (0.27)	1.25	0.89** (0.21)	2.43	0.11 (0.32)	1.11	0.67** (0.25)	1.96
HumRghts	--	--	--	--	0.93** (0.16)	2.53	0.61** (0.10)	1.83
DissViol	0.09** (0.04)	1.09	0.05* (0.03)	1.05	0.08** (0.05)	1.09	0.02 (0.03)	1.02
Civil War	3.25** (0.24)	25.9	1.04** (0.18)	2.85	2.44** (0.28)	11.5	0.60** (0.20)	1.82
War on Ter	1.47** (0.52)	4.35	0.47 (0.39)	1.60	1.31** (0.62)	3.70	0.23 (0.47)	1.26
Democ	0.04** (0.02)	1.04	-0.06** (0.01)	0.94	0.06** (0.02)	1.06	-0.02** (0.01)	0.98
Trans	-0.09 (0.33)	0.91	0.34 (0.27)	1.40	-0.43 (0.38)	0.65	-0.03 (0.30)	0.97
GNP	-5.45^{e-12}# (2.06^{e-12})	1.00	-9.41^{e-13}** (5.08^{e-13})	1.00	-7.60^{e-12}# (2.52^{e-12})	1.00	-4.54^{e-13} (4.41^{e-5})	1.00
NbhdGen	0.57 (0.99)	1.77	-1.76** (0.71)	0.17	1.12 (1.37)	3.06	-1.17* (0.91)	0.31
NbhdHumRts	--	--	--	--	0.44 (0.56)	1.55	0.22 (0.32)	1.25
NbhdDissVi	-0.48 (0.50)	0.62	-0.53** (0.29)	0.59	0.26 (0.60)	1.29	-0.22 (0.35)	0.80
NbhdGNP	0.12 (0.38)	1.13	0.33* (0.22)	1.39	-0.20 (0.44)	0.82	0.56** (0.26)	1.75
NbhdDemoc	0.91# (0.40)	2.49	1.28** (0.23)	3.61	0.31 (0.45)	1.37	1.15** (0.26)	3.15
NbhdCivWar	3.07** (0.65)	21.6	1.57# (0.44)	4.80	1.67** (0.79)	5.33	0.01 (0.55)	1.01
NbdWoT	-5.00 (3.00)	0.01	-0.13 (1.07)	0.88	-6.48 (5.36)	0.00	0.65 (1.20)	1.91
Mountains	0.19 (0.51)	1.20	0.53# (0.26)	1.70	0.10 (0.59)	1.11	0.94# (0.34)	2.56
$ForceStock_{t-1}$	5.77^{e-7}** (1.39^{e-7})	1.00	4.86^{e-7}** (1.14^{e-7})	1.00	2.06^{e-7}* (1.47^{e-7})	1.00	1.53^{e-7} (1.31^{e-7})	1.00
Constant	-4.71** (0.41)	-	-3.23** (0.23)	-	-6.67** (0.75)	-	-4.81** (0.46)	-
N	3,218				2,171			

** Significant at the .05 level (one-tailed test)
* Significant at the .10 level (one-tailed test)
# Significant at the .10 level, but in the wrong direction (two-tailed test)

Table 2: Variables with Differential Effects

	IDPs > Refugees	**Refugees > IDPs**
Local	Democracy (+) Civil War (+) War on Territory (+)	Democracy (-) Civil War (+) Genocide (+)
Neighborhood	--	Democracy(+) Genocide (-) GNP (+)