

## Letter

# The Causes and Consequences of Refugee Flows: A Contemporary Reanalysis

ANDREW SHAVER *University of California, Merced, United States*

BENJAMIN KRICK *Duke University, United States*

JUDY BLANCAFLOR *Political Violence Lab, United States*

XAVIER LIU *Political Violence Lab, United States*

GHASSAN SAMARA *Political Violence Lab, United States*

SARAH YEIN KU *University of Michigan, United States*

SHENGKUO HU *Independent Researcher, United States*

JOSHUA ANGELO *Political Violence Lab, United States*

MARTHA CARREON *Political Violence Lab, United States*

TRISHIA LIM *Political Violence Lab, United States*

RACHEL RAPS *Political Violence Lab, United States*

ALYSSA VELASQUEZ *Political Violence Lab, United States*

SOFIA DE MELO *Political Violence Lab, United States*

ZHANYI ZUO *Political Violence Lab, United States*

**T**he world faces a forced displacement crisis. Tens of millions of individuals have been forced across international boundaries worldwide. Therefore, the causes and consequences of refugee flows are the subjects of significant social science inquiry. Unfortunately, the historical lack of reliable data on actual refugee flows, country-specific data reporting timelines, and more general pre-2000 data quality issues have significantly limited empirical inferences on these topics. We replicate 28 articles on these topics using data newly released after a multiyear collaboration with the United Nations on annual dyadic flows. We observe major inconsistencies between the newly released flow numbers and the stock-based flow estimates upon which decades of research are based; we also find widespread inappropriate treatment of missing historical values. When we replicate the existing literature using the newly introduced flow data, correcting the treatment of missing historical values, and temporally extending/restricting the study periods, we produce significantly different results.

Corresponding author: Andrew Shaver , Assistant Professor, Department of Political Science, University of California, Merced, United States, [ashaver@ucmerced.edu](mailto:ashaver@ucmerced.edu).

Benjamin Krick , Doctoral Student, Department of Political Science, Duke University, United States, [benjamin.krick@duke.edu](mailto:benjamin.krick@duke.edu).

Judy Blancaflor, Research Intern, Political Violence Lab, United States, [judy.blanc99@gmail.com](mailto:judy.blanc99@gmail.com).

Xavier Liu, Research Intern, Political Violence Lab, United States, [xavier.liu17@gmail.com](mailto:xavier.liu17@gmail.com).

Ghassan Samara, Research Intern, Political Violence Lab, United States, [gsamara@ucmerced.edu](mailto:gsamara@ucmerced.edu).

Sarah Yein Ku, PhD Student, Department of Political Science, University of Michigan, United States, [sarahyku@umich.edu](mailto:sarahyku@umich.edu).

Shengkuo Hu, Independent Researcher, United States, [shu8@ucmerced.edu](mailto:shu8@ucmerced.edu).

Joshua Angelo, Research Intern, Political Violence Lab, United States, [joshua.angelo@duke.edu](mailto:joshua.angelo@duke.edu).

Martha Carreon, Research Intern, Political Violence Lab, United States, [martha.carreon.100@my.csun.edu](mailto:martha.carreon.100@my.csun.edu).

Trishia Lim, Research Intern, Political Violence Lab, United States, [trishia\\_lim@berkeley.edu](mailto:trishia_lim@berkeley.edu).

Rachel Raps, Research Intern, Political Violence Lab, United States, [rraps@berkeley.edu](mailto:rraps@berkeley.edu).

Alyssa Velasquez, Research Intern, Political Violence Lab, United States, [aavelas2@uci.edu](mailto:aavelas2@uci.edu).

Sofia de Melo, Research Intern, Political Violence Lab, United States, [sofiademelo98@gmail.com](mailto:sofiademelo98@gmail.com).

Zhanyi Zuo, Research Intern, Political Violence Lab, United States, [zzy19961013@berkeley.edu](mailto:zzy19961013@berkeley.edu).

Received: September 06, 2022; revised: March 29, 2023; accepted: February 27, 2024. First published online: May 07, 2024.

## INTRODUCTION

**T**he world faces a forcible displacement crisis: more than one hundred million individuals have either fled their countries or been displaced within them.<sup>1</sup> International displacement has risen dramatically in recent decades due to conflict and political instability in countries from Afghanistan to Venezuela, with Ukraine producing the single largest outflow of refugees in a single year in recorded history.<sup>2</sup> Forcible displacement has tremendous human costs, and its causes and consequences are the focus of significant social science research. However, empirical studies of refugee flows have been limited by a lack of reliable data.

Existing results are often derived from refugee population stock measures—total end-of-year population counts of refugees within host countries—rather than actual flow estimates. Stock-based results are subject to fundamental questions about internal validity. Much prior research is based on data collected before 2000, when the United Nations Refugee Agency (UNHCR) began systematically tracking refugee and asylum seeker (REF/ASY)<sup>3</sup> numbers globally. The quality of pre-2000 data is limited, with many missing values that much existing research does not properly account for. Analyses focused on origin countries are often missing data because asylum countries were not then reporting arrivals to the UNHCR; separately, asylum countries that were reporting arrival data often did not collect national origin.

In this study, we seek to address these issues. Following a multiyear collaboration with the UNHCR, culminating in the release of new international displacement flow data (1962–2022),<sup>4</sup> we reevaluate 28 studies published over decades on the causes and consequences of refugee flows.<sup>5</sup> Using country-specific reporting timelines, we update these articles' results to account for possible measurement error introduced when missing values were treated as 0s. We also temporally extend these studies so that results are based on contemporary observations less affected by historical data-quality issues. Our goal is to understand how existing results change when we address these issues—an important question for a body of work with significant influence on academic and policy discourses.

In brief, we observe large inconsistencies between the newly released flow numbers and the stock-based estimates upon which decades of research is based, and

we find that the inappropriate treatment of missing historical values is widespread. We produce significantly different results when we replicate the existing literature following three different approaches: first, replacing the old stock-based data with the newly introduced flow data; second, correcting the treatment of missing historical values; and third, temporally extending the studies.

Specifically, we find that in 19 articles on flow causes,  $\approx 74\%$  of findings replicate; in 9 articles focused on flow consequences, only  $\approx 50\%$  replicate. These percentages are conservative: we assess only whether previously reported results maintain statistical significance and/or whether the sign of estimated coefficients reverses. A stricter standard would also assess substantial changes in magnitude (toward 0), likely driving these percentages lower still.

A subset of our replications are “theory-based”—these consist of reanalyses of studies that focus on refugee flows but adopt some other measure of REF/ASY (e.g., stocks).<sup>6</sup> These studies contribute substantially to the ongoing debate about the effects of refugees on political violence in host countries—as Savun and Gineste (2019, 88) note, “security consequences associated with refugee flows are among the most widely studied aspects of forced migration.” Thus, we view these replications as complements to the original work, offering new empirical insights, and we clearly distinguish in the results which replications are theory-based. In our replications, effects of refugees on security conditions are attenuated, suggesting that the literature's identification of refugees as sources of violent instability is likely overstated. The contrast between our findings and those based on stocks points to potentially important differences in the effects of refugee inflows versus sustained presence.

The new data also reveal that forced displacement is much more common than reported: Rubin and Moore (2007, 91) note that “[f]orced migration is a relatively rare event ... around 82% [of country-year cases] experienced no forced migration ... .” When we extend this study to 2000–21, we calculate that number to be 22%.<sup>7</sup>

## ON THE INTERNAL VALIDITY OF PREVIOUS RESULTS

### Refugee Measurement

The UNHCR has tracked REF/ASY flows since 1962. Flow records were used primarily for operational purposes and were not centralized until recently. In 2019, the UNHCR released a draft flow dataset. We engaged in extensive discussions with UNHCR staff about the data, including possible additions/modifications to capture new international movements and apparent incon-

<sup>1</sup> See <https://www.unhcr.org/global-trends>.

<sup>2</sup> We later discuss the uncertainty surrounding the magnitude of these increases due to the constraints of available data.

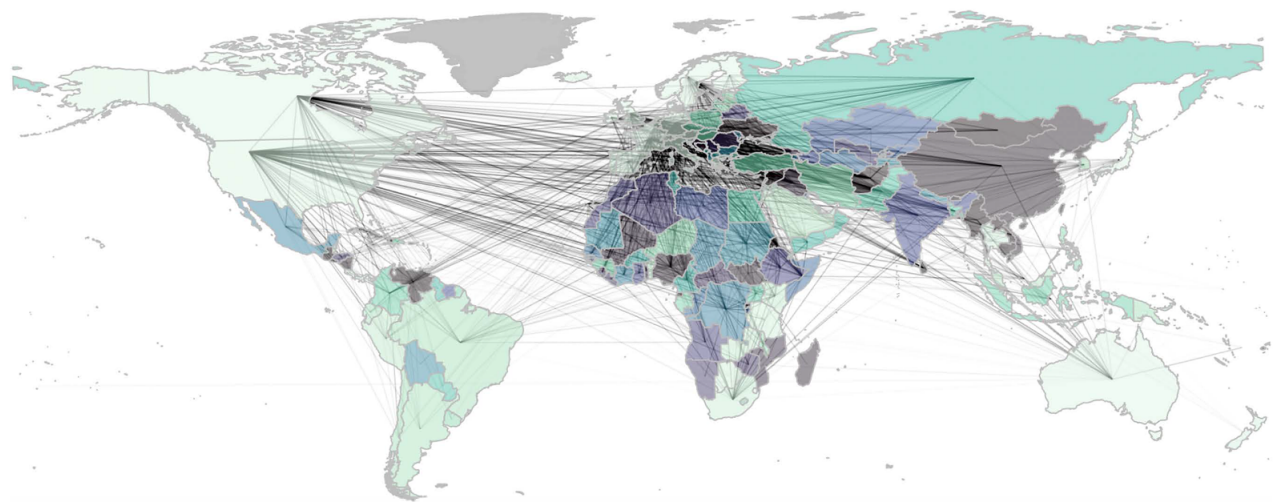
<sup>3</sup> In some cases, refugees and asylum seekers are distinct populations. In others, asylum applications are simply the mechanism by which individuals are subsequently registered as refugees. See the Supplementary Material for additional discussion.

<sup>4</sup> For details, see <https://www.politicalviolencelab.com/refugeeflows>.

<sup>5</sup> Specifically, we reassess the causes of REF/ASY outflows from origin countries and the impacts of REF/ASY inflows into asylum countries. “Flows” may refer to the number of departures from an origin country (“outflows”), arrivals to an asylum country (“inflows”), or individuals fleeing one country for another (“directed dyadic flows”). When referring to these movements generally, we adopt the term “flow(s),” specifying specific types as relevant to specific research questions, data constructions, etc.

<sup>6</sup> For example, authors avoiding using stock-based flow estimates because of data quality concerns or using stock figures after theorizing about the effects of both flows and stocks.

<sup>7</sup> For asylum country-years without (in)flows, that percentage is 28%.

**FIGURE 1. Global Dyad-Year Refugee and Asylum Seeker Outflows**

*Note:* Dyad-year refugee and asylum seeker outflows (between country centroids for all outflows greater than or equal to five hundred persons for the period of 2000–21) using the new UNHCR data. Country color coding reflects percentage of flows into and out of each country's borders that are outflows.

sistencies across data versions.<sup>8</sup> We also compared statistical results using redacted and unredacted data versions to help validate the UNHCR's decision to release only redacted data to protect individual asylum seekers' identities in cases of very small dyadic flows.

The UNHCR ultimately released the final "Forced Displacement Flow Dataset" in 2022. The new flow data are depicted in Figure 1, with additional details in Appendix A.1.1 of the Supplementary Material.

#### *Flows and Stock-Based Estimates Compared*

Actual flows diverge from stock-based estimates for several reasons.<sup>9</sup> First, researchers estimate flows from stocks as follows, where  $i$  denotes either the directed dyad (i.e., sending–receiving country pair) or the asylum/origin country, depending on the unit of interest, and  $t$  refers to the year:

$$\hat{f}_{i,t} = \Delta s_{i,t} = \begin{cases} s_{i,t} - s_{i,t-1}, & \text{when } s_{i,t} \geq s_{i,t-1}, \\ 0, & \text{when } s_{i,t} < s_{i,t-1}. \end{cases} \quad (1)$$

Stock-based estimates calculated this way do not account for naturalizations, returns, or resettlements (hereafter "stock departures"); births and deaths; or any other variable affecting host-country stock levels. We find that a substantial number of (directed-dyad) cases involve simultaneous (same-year) stock departures and directed flows. In  $\approx 45.40\%$  of asylum country-year observations, inflows and stock departures co-occur; in  $\approx 20.81\%$  of cases, stock departures are greater

than or equal to inflows (see Figure 2). By capturing new arrivals, the new flow data avoid this issue.

Second, under the stock-based estimation approach, years of "negative" flow are set to 0. We calculate that nearly one-third ( $\approx 29.68\%$ ) of all first-differenced observations result in negative values that are converted to 0s. In just under half of these directed-dyad-year cases ( $\approx 48.56\%$ ), the new data report positive values instead.

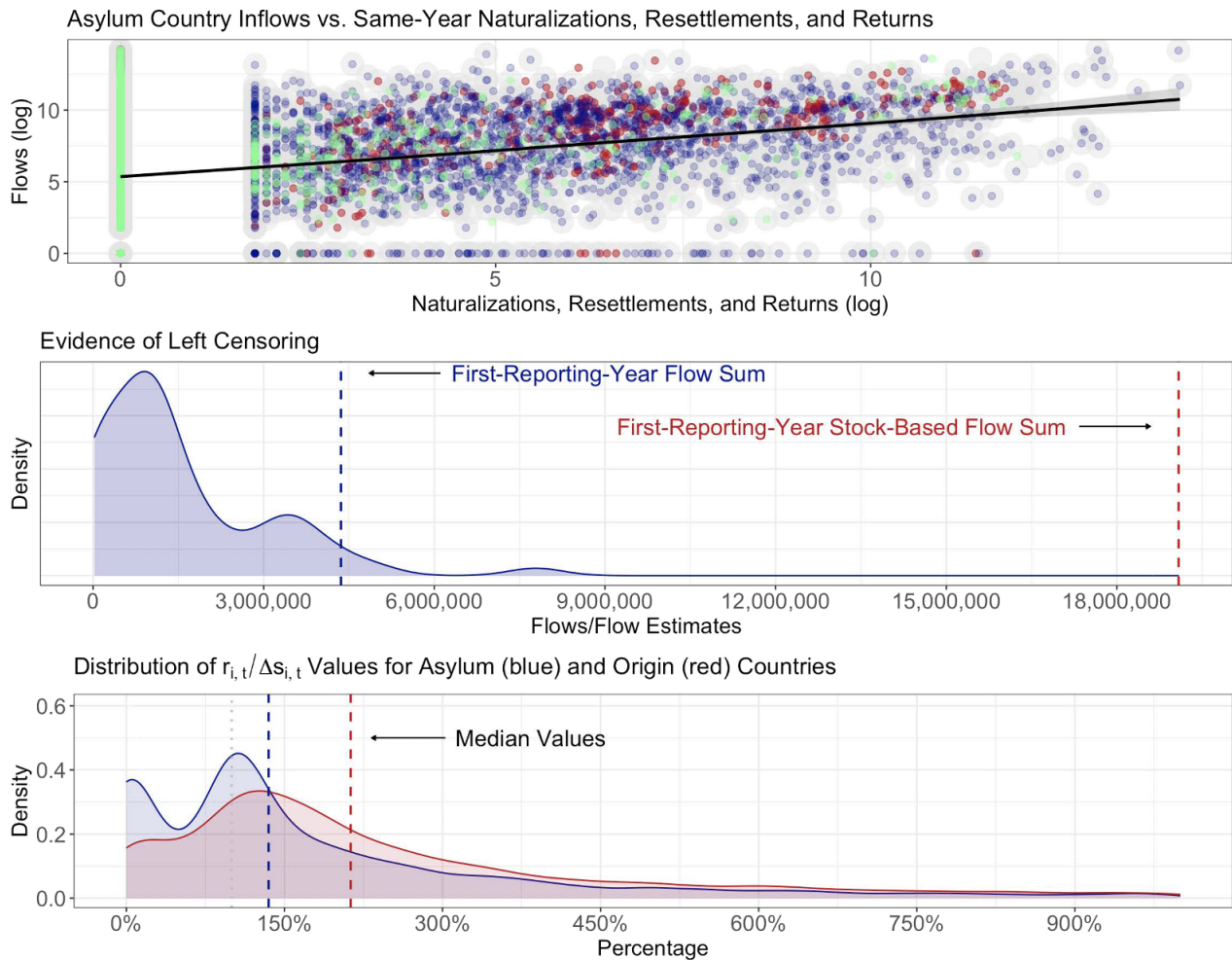
Third, stock-based flow estimates suffer from major left-censoring. The UNHCR begins tracking REF/ASY arrivals for different countries in different years. Under the first-differences approach, estimates may capture preexisting populations (not inflows) for the first year in which a positive value is reported. To quantify this potential issue, we compare the sum of refugee stocks for all directed-dyad-year observations corresponding to the first year of UNHCR reporting to the sum of the new data's flows for those same years. Results suggest that many stock-based flow estimates capture preexisting refugee populations rather than new flows—a source of significant potential error in statistical estimates (see Figure 2). Preexisting population values do not enter into the new flow data.

Fourth, until 2007, stock data include population values for third country resettlements, erroneously depicting "flows" into countries where REF/ASY eventually resettled, sometimes years after displacement. The new data prioritize asylum seeker applications to reflect increases in the year of their actual arrival.

Fifth, the stock data include "non-flow increases": adjustments to stock values due to methodological revision, legislative change, or other host-country changes to how REF/ASY are defined or calculated. These positive reestimations produce apparent flow increases that do not reflect actual new arrivals. In

<sup>8</sup> These discussions also involved Fearon and Shaver (2020).

<sup>9</sup> For a more detailed discussion, see Appendix A.1.2 of the Supplementary Material.

**FIGURE 2. Measurement Issues Associated with Stock-Based Flow Data**

*Note:* From top to bottom: Plot 1: This figure plots asylum country-year refugee naturalizations (green), resettlements (red), and returns (blue) and against inflows. Broader gray dots represent the sum of naturalizations, resettlements, and returns for each asylum country-year. The figure shows that significant numbers of refugees/asylum seekers are often naturalized, resettled, and/or returned in the same years that refugees continue to arrive. In such years, stock-based inflow estimates will be skewed downward. Data on naturalizations, resettlements, returns, and inflows provided by the UNHCR (UNHCR 2021a, 2021b). Plot 2: This figure provides strong evidence of a left-censoring effect in stock-based flow estimates. Specifically, inflows estimated using stocks show significant spikes on the first year of UNHCR country reporting that likely reflect preexisting refugee populations, not actual new inflows. Plot 3: This figure displays estimated “bias” in stock-based estimates of inflows given by  $\frac{\text{inflow}_{i,t}}{(\text{stock}_{i,t} - \text{stock}_{i,t-1})}$ . The figure displays the distribution of resulting percentages for (a) all asylum-country year observations and (b) all origin-country year observations. Overall, these percentages fall well above 100%, indicating that stock-based estimates generally significantly underestimate actual inflows.

the new flow data, these changes have been removed.<sup>10</sup>

Sixth, stock-based flows lag actual flows in countries that use their asylum systems to grant refugee status; asylum seekers enter into stock data only after their asylum applications have been processed and approved—sometimes years after arrival.<sup>11</sup> The new flow data

prioritize asylum seeker *applications* to capture movements during the years in which they occurred.

Overall, how do the new flows compare with stock-based estimates?  $s_{i,t}$  is a function of stock departures. When stock departures occur simultaneously (within the same year) with inflows, measures of flows are attenuated:  $\hat{f}_{i,t} \leq f_{i,t}$  (i.e., the number of stock departures in a given year reduces the calculable inflows by that number). This is consistent with patterns in the data: in  $\approx 81\%$  of directed-dyad-year cases, flow values are strictly larger than stock-based estimates (and larger than or equal in  $\approx 84\%$  of cases). Overall, we

<sup>10</sup> There is one important caveat (see Appendix A.1.5 of the Supplementary Material for details).

<sup>11</sup> See Appendix A.1.1 of the Supplementary Material for an important discussion on non-approved asylum applications.

calculate that the new flow data capture 14,227,372 more flows than the stock-based data from 1962 to 2022<sup>12</sup>; for every  $\approx 5$  flows reported under the stock-based approach, the new data report one additional flow.

We directly compare flow values with their stock-based estimates, estimating bias as  $\frac{\text{inflow}_{i,t}}{(\text{stock}_{i,t} - \text{stock}_{i,t-1})}$ . Figure 2 displays the distribution of resulting percentages for (a) all asylum-country year observations and (b) all origin-country year observations. Overall, these percentages fall above 100%.<sup>13,14</sup> In Appendix A.1.3 of the Supplementary Material, we supplement this analysis by reporting for each asylum country the correlation between inflows to that country and stock-based estimates. Results indicate that stock-based estimates tend to significantly underestimate flows; in  $> 10\%$  of cases, the two variables are either not correlated or are negatively correlated.

### *Pre-2000 Data Missingness and Quality Issues*

Three major issues are associated with UNHCR data generation and reporting patterns before the year 2000, when the UNHCR standardized approaches to data collection and when many asylum states adopted information and communication technologies that significantly improved reporting.<sup>15</sup> The empirical problems discussed below persist beyond the year 2000, but are significantly reduced; we use the pre-/post-2000 framing for analytical parsimony.<sup>16</sup>

The first empirical issue is the inappropriate treatment of missing data. Until recently, centralized data on when the UNHCR began tracking REF/ASY in each country were unavailable. In the absence of positive displacement values, many panel datasets set country-/dyad-year observations to 0.<sup>17,18</sup> While some missing positive refugee values for yearly country-/dyadic observations may reflect true 0s, others still reflect positive values that the UNHCR did not collect. Nearly every study we replicated follows the practice of setting such observations to 0 when they precede country-specific data collection timelines. For an asylum-country-year panel dataset 1962–99, this

practice results in  $\approx 49.82\%$  of observations being set to 0.<sup>19</sup>

We supplement our analysis with UNHCR-supplied data on centralized collection efforts by country from 1970 on.<sup>20</sup> Patterns in data collection are depicted in Figure 3. Many countries' data do not appear in centralized records until long after statistics began to be collected. Before 2000, UNHCR collected asylum seeker data only from several dozen industrialized countries; in 2000, when they centralized data collection, that number jumped to 137 countries, with more countries being added every year.

Using the new flow data, we produce panel datasets with observations set to NA (rather than 0) for years before data collection began.<sup>21</sup> As we show in Appendix A.3.1 of the Supplementary Material (and in the full set of results posted in a secondary appendix in the Dataverse; see Shaver et al. 2024), this replacement produces additional changes in several results.

The second empirical issue emerges because UNHCR records are mostly constructed from asylum state records: studies using origin-country panel datasets are missing some unknown (potentially very substantial) number of REF/ASY outflows. These missing values were not captured by corresponding inflow data from asylum countries that were not yet reporting data to the UNHCR (see Figure 3). Approximately 68% of the “causes” studies we replicate (and  $\approx 40\%$  of all of the studies we replicate) use origin-country panel datasets.

The third empirical issue is that until 2000, a significant amount of UNHCR data for tracked REF/ASY are missing national origin information. For research designs in which REF/ASY origin is relevant, missingness on this variable introduces noise (and potentially bias) to results. We display this pattern in Figure 3.

We cannot directly correct for these final two issues. However, by 2000, these problems are substantially eliminated. For this reason, our analyses include contemporary replications: we extend studies through the most recent date possible and analyze them from the period beginning in 2000. These are our preferred specifications, as all four issues that we raise are either resolved or substantially mitigated.<sup>22</sup>

<sup>12</sup> This figure excludes stateless and Palestinian refugees from both the new flow and stock datasets (see Appendix A.1.1 of the Supplementary Material).

<sup>13</sup> Outlying values are produced when flows significantly exceed the stock-based estimates, with resulting means exceeding 1,000%; we therefore use median values (134.92% and 212.84% for the asylum and origin countries, respectively).

<sup>14</sup> This approach drops observations in cases where  $\Delta\text{stock} < 0$  and values are set to 0. At the asylum- and origin-country levels, these percentages are 15.13% and 24.83%, respectively.

<sup>15</sup> See Marbach (2018) for additional discussion of these issues and a proposed solution.

<sup>16</sup> See Appendix A.1.4 of the Supplementary Material for a more detailed treatment.

<sup>17</sup> Typically, authors constructed balanced panels, setting country-/dyad-year observations to the same starting year and assigning 0s to observations without refugee values.

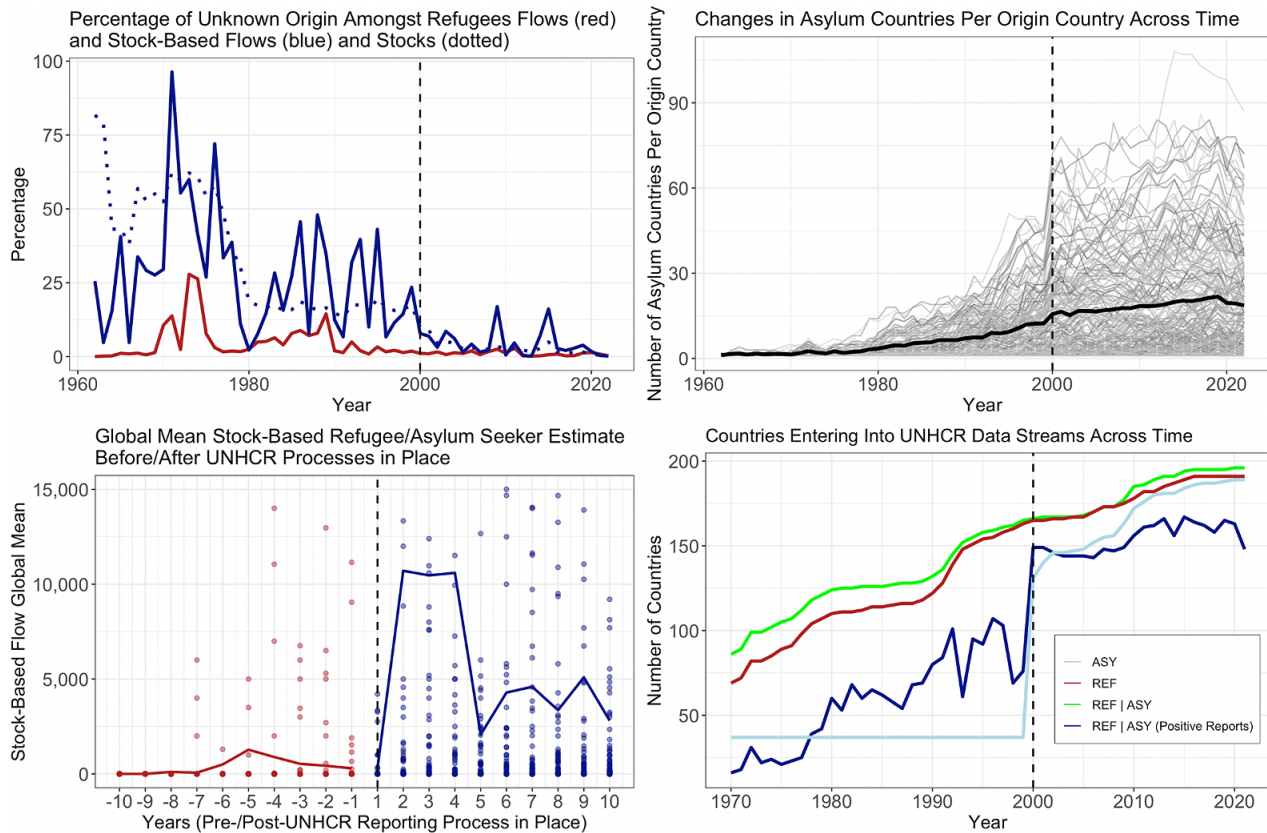
<sup>18</sup> This practice is widespread. Of the 28 articles we analyze, 24 impute 0s; 4 avoided this issue by limiting their analyses to modern periods.

<sup>19</sup> For a directed-dyad panel, the percentage is  $\approx 50.75\%$ .

<sup>20</sup> The exception is the set of 37 countries that supplied data to the UNHCR on asylum seeker flows 1970 to 1999. For these countries, initial reporting years are unknown; we know only that these 37 countries reported asylum figures for some or all years during this period. We discuss this in more detail later.

<sup>21</sup> As discussed below, data on flows were secured at the level of the asylum country (not the origin country). We can identify only a subset of observations with missing data in our origin country-year panel dataset. We therefore construct origin-country panel datasets using a more complicated procedure, described in Appendix A.2.4 of the Supplementary Material.

<sup>22</sup> Data issues aside, there are other reasons why contemporary study results might differ from previous results, ranging from overall displacement numbers to the evolution of the international response to flows.

**FIGURE 3. Temporal Changes in UNHCR Data Collection and Quality**

*Note:* Clockwise from top left: Plot 1: This figure displays the percentage of directed-dyad-year observations in which the origin of the refugee/asylum seeker population is unknown in the new flow data (red), the stock-based estimates (blue) and stocks (dotted). In many instances, the national origin of refugees/asylum seekers was not tracked; the number of these instances fall drastically by 2000. Plot 2: This figure displays the annual number of asylum countries reporting outflows for each origin country, which significantly increases by 2000. Pre-2000 origin-country data are very likely missing many outflows from those countries, for many of the countries to which they fled did not report them. Plot 3: This figure plots (in green) the number of countries reporting refugee or asylum seeker inflow data to the UNHCR. Lines in red and light blue disaggregate yearly country totals between refugees and asylum seekers, respectively. The darker blue depicts the total number of countries per year for which actual refugee or asylum seeker numbers were reported. Differences in the green and blue lines may reflect cases in which potential asylum states had data sharing agreements in place with the UNHCR but did not have any numbers to actually report. Plot 4: This figure displays stock-based inflow estimates for each asylum country for the decade preceding and following the year in which a UNHCR reporting process was put in place in that country. Mean stock-based inflow estimates for all asylum countries are plotted for each year. Points depict individual asylum-country values. If post-reporting process trends generally reflect actual pre-reporting process trends, then pre-reporting values adopted by scholars (virtually all 0s) are likely systematically skewed toward 0.

## REPLICATIONS

We used Google Scholar to search general and social science academic journals for articles engaging in quantitative research on global refugee flows. Our search query limited cases to those that (i) reference refugee flows, (ii) include the terms “UNHCR” and “data,”<sup>23</sup> and (iii) were published by a major publisher,<sup>24</sup> returning 1,556 responsive articles. We manually inspected each, eliminating studies that (i) did not deal with causes or consequences of refugee flows, (ii) were

entirely qualitative, (iii) incorporated data on refugees only as a control or in secondary (tertiary, etc.) analyses, or (iv) were single country studies. This produced 35 qualifying studies. We were unable to obtain replication materials for seven of these. We replicated the remaining 28 by (i) correcting incorrectly imputed zeros and (ii) replacing the old stock-based measures with the new flow data. We assess whether previously reported results maintain statistical significance and/or whether the sign of estimated coefficients reverses.<sup>25</sup> A detailed description of these articles appears in Appendix A.2.2 of the Supplementary Material; more

<sup>23</sup> Because data on refugee stocks (whether used to estimate flows or used directly) are supplied by the UNHCR, this term helps restrict results to quantitative studies.

<sup>24</sup> See Appendix A.2.1 of the Supplementary Material for the list of publishers and formal search query.

<sup>25</sup> A stricter standard would also assess changes in magnitude. Accordingly, our replication standard is conservative.

information on replication procedures is included in Appendix A.2.4 of the Supplementary Material.

By assessing how existing results change when we address the empirical issues described above, this effort falls into the class of “broad” (Dafoe 2014), “statistical” (Hamermesh 2007), or “wide” (Pesaran 2003) replications involving reestimating test results with the use of new data or related modifications (e.g., adopting alternatively constructed variables). The studies we replicate form the backbone of research on the causes and consequences of refugee flows and have influenced research agendas, curricula, and policy discourses.<sup>26</sup> Replication results inform causal inferences in cases in which the original authors’ testing strategies were well-identified—save for the empirical corrections we apply—but more generally, they update our understanding of the “published record ... recognized [as] state of the art” (King 2006, 119), providing direction for additional scholarly inquiry and the reexamination of their policy implications.<sup>27</sup>

## RESULTS AND DISCUSSION

Results are succinctly presented in Tables A1 and A2 in Appendix A.3.1 of the Supplementary Material. Complete replication regression results (alongside original estimates) appear in a supplementary appendix posted to Dataverse (Shaver et al. 2024). Of the 62 total tests from causes articles,  $\approx 74\%$  replicate. More significantly, of 20 total tests from consequences articles, only  $\approx 50\%$  replicate.<sup>28</sup> We classify 14 of the 28 (50%) articles we replicate as “plausibly causally identified.”<sup>29</sup> Of these,  $\approx 69\%$  of results on causes replicate and  $\approx 50\%$  focused on consequences replicate. This is quantitatively and substantively consistent with the results in our full sample. We present and discuss a select set of results below.

<sup>26</sup> This research totals 4,257 citations at the time of writing.

<sup>27</sup> Our research follows other replication efforts of this nature (e.g., Carroll and Kenkel 2019; Fisher, Johnson and Smeeding 2013; Nystrup and Bramwell 2020), though we note that there is no clearly established, single definition of, or approach to, replication in the social sciences (Clemens 2017; Duvendack, Palmer-Jones, and Reed 2017; Freese and Peterson 2017). Consistent with the practical research realities inherent to global, quantitative refugee research in which observational studies are prevalent, our goal is not to produce a causally interpretable meta-analysis more typical of academic disciplines wherein causally identified research designs (e.g., randomized controlled trials) are more practical and common. Instead, our focus on fundamental data issues seeks to facilitate continued causal inferential progress within the forcible displacement research space.

<sup>28</sup> This is partially driven by “theory based extensions” (see Appendix A.2.3 of the Supplementary Material), and these results thus illuminate the difference between the effect of refugee flows versus refugees’ continued presence.

<sup>29</sup> This classification is based on the use of instrumental variable, difference-in-difference, regression discontinuity, synthetic control, natural experiment, or fixed effects study designs (see Supplementary Tables A1 and A2). We adopt this criterion following referee comments.

## Causes

With respect to the causes of flows, updated study results rarely overturned original findings; however, they frequently supported hypotheses discarded by the original authors as statistically unsupported.

Our findings confirm the central roles played by the “push factors” of political violence and state repression in driving international displacement. We corroborate results that link civil war/insurgency to outflows, estimating larger effects of these factors than did Echevarria and Gardeazabal (2021), Davenport, Moore, and Poe (2003), and Moore and Shellman (2004); we also uncover a larger effect of state repression on outflows than Rubin and Moore (2007). Steele (2017, 9) has observed that “current understanding tends to equate wars or violence with an increase in displacement, but we can and need to be more precise.” Our replications amplify her call. Future work might incorporate the new flow data into global analyses of potential heterogeneous effects across factors such as the timing of violence, its spatial distribution and intensity, and the technologies used to perpetrate it.

We also replicated papers focused on “pull factors” that incentivize international over internal displacement and influence the choice of international destination. In replications of Moore and Shellman (2004; 2007) and Turkoglu and Chadeaux (2019), we find limited support for the idea that refugees are motivated by economic opportunity or democratic institutions in destination countries. This finding contrasts with the framing of asylum seekers as opportunists—as echoed in prominent political discourse. Other results raise tensions warranting further study: for instance, regarding the role of alliance dynamics, we fail to substantiate Moorthy and Brathwaite’s (2019) finding that the presence of formal alliances positively influences dyadic flows. However, whereas Moore and Shellman (2007) do not find this, we do.

A more subtle theme of our replications is the underexplored role of factors discouraging or restricting individuals from seeking refuge abroad. On the one hand, some updated findings point to the role of restrictions. In our replication of Echevarria and Gardeazabal (2021), we estimate larger effects of country size, proximity to potential asylum states, and island status. We estimate a larger effect size than Moore and Shellman (2007) of potential asylum state contiguity. On the other hand, we find little connection in the Moore and Shellman (2007) replication between conflict and repression in potential asylum states and refugee inflows. These and other such findings encourage further broader inquiry into the set of factors responsible for restricting international displacement—from border securitization (Simmons and Kenwick 2022) to severe weather and natural disasters along border regions under climate change.

## Consequences

With respect to the consequences of inflows, we observe significant changes from previous results. In

our replications of this seminal literature, the relationship between the arrival of refugees and the onset of war and political violence is attenuated: we find that refugees are only infrequently conduits of violence, and the conditions under which forced displacement poses a risk to host countries appear to be specific.

With respect to refugees' connection to terrorism, our replications find only partial support for Choi and Salehyan's (2013) analysis linking these variables. This is consistent with Milton, Spencer and Findley's (2013) results and our corresponding replication (though we estimate smaller effect sizes). We corroborate Polo and Wucherpfennig's (2022) causal finding that refugee influx is positively associated with terrorism in the specific case of refugees from communities with ties to transnational terrorist organizations; we find additional evidence that the association for refugees originating from countries without ties is *negative*. Findings highlight the potential heterogeneous treatment effects of inflows on terrorism, with potential implications for more tailored policy responses and programming.

When we reexamine work on refugees and governments' respect for human rights, we fail to confirm either of Wright and Moorthy's (2018) findings: our results indicate neither that an influx of refugees positively influences repression nor that this relationship is moderated by development. We do not recover Chu's (2020) findings relating to refugees from rival and non-rival origin states and hosts' respect for human rights.

Finally, when we replicate work linking refugees to inter- and intrastate conflict, our findings do not substantiate Salehyan and Gleditsch's (2006) seminal research associating refugees to civil war diffusion.<sup>30</sup> Our results partially support Salehyan's (2008) findings that flows between states can provoke militarized disputes: flows between a given dyad increase the probability that the receiving state initiates a conflict with the sender, but we do not find that it increases the probability of sender-state initiation.

Collectively, these findings speak against the new politics of fear, challenging political narratives that frame refugees as security threats and the restrictive state policies they underpin. Our findings do not indicate that there are *no* effects of refugee inflows on violent instability, but it seems that refugees play a role in producing or facilitating political violence, wittingly or not, only under particular circumstances. The differences between our findings and those based on refugee stocks highlight potentially important differences between the effects of refugee inflows into a country and the effects of sustained refugee presence. These differences warrant further exploration—particularly because of the “growing difficulty [of] uprooted people ... in finding lasting solutions to their plight” (Crisp 2021, 3).

We conclude by noting a publication bias against null results (Esarey and Wu 2016; Gerber and Malhotra 2008). Our replications sometimes supported hypotheses that were discarded when tested with lower-quality data because they lacked statistical support. Other meaningful relationships relating to refugee flows may have therefore gone undiscovered, which scholars might now retest with these new data.

## SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055424000285>.

## DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/JADOZL>. Limitations on data availability are discussed in the text and/or Supplementary Material.

## AUTHORS CONTRIBUTIONS

The first two authors are the primary authors and are listed in order of contribution. The remaining authors are (former) Political Violence Lab members, listed in order of their respective contributions (and alphabetically where contributions were equal).

## ACKNOWLEDGEMENTS

We are grateful for the opportunity to work with the United Nations Refugee Agency (UNHCR) on its release of the data that enabled this project. We thank the University of California, Merced's University Library Collection Services for supporting the acquisition of data for this project. The Political Violence Lab, based at the University of California, Merced, acknowledges the support it has received from the University of California, Merced and the University of California Washington Center. We further thank the *American Political Science Review* editorial team and the three anonymous referees, Lamis Abdelaaty, Kyle Beardsley, Mietek Boduszyński, Alex Bollfrass, Alex Braithwaite, Mateo Villamizar Chaparro, Elaine Denny, James Fearon, Guy Grossman, Biz Herman, Connor Huff, Adam Lichtenheld, Bryce Loidolt, Eric Mvukiyehe, Fouad Pervez, Ryan Powers, David Siegel, Abbey Steele, Yang-Yang Zhou, and participants of Duke University's Political Economy Seminar Series; Harvard University's Political Violence Workshop; the Annual Peace Science Society International Conference; and the Graduate Students in International Relations etc. (GSISE) Seminar for comments on this project. For research assistance, we gratefully acknowledge Political Violence Lab interns Mairead Allen, Mia

<sup>30</sup> See additional details on this complicated case in the Supplementary Material with results posted to Dataverse (Shaver et al. 2024); the relationship originally reported may have been driven by post-World War II migration dynamics and/or limited data from this period. We invite additional investigation.

Bartschi, Kai Walter Bauer-Seeley, Isabella Caldarelli, Dheera Dusanapudi, Jasmine Her, Zainab Khan, Qi Zhi (Aaron) Liow, Audrey Lozano, Yasmine Lunar, Colby Mathe, Brian McCarthy, Daniel Romero, and Danielle Startsev.

## CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

## ETHICAL STANDARDS

The authors affirm this research did not involve human participants.

## REFERENCES

- Carroll, Robert J., and Brenton Kenkel. 2019. "Prediction, Proxies, and Power." *American Journal of Political Science* 63 (3): 577–93.
- Choi, Seung-Whan, and Idean Salehyan. 2013. "No Good Deed Goes Unpunished: Refugees, Humanitarian Aid, and Terrorism." *Conflict Management and Peace Science* 30 (1): 53–75.
- Chu, Tiffany S. 2020. "Hosting Your Enemy: Accepting Refugees from a Rival State and Respect for Human Rights." *Journal of Global Security Studies* 5 (1): 4–24.
- Clemens, Michael A. 2017. "The Meaning of Failed Replications: A Review and Proposal." *Journal of Economic Surveys* 31 (1): 326–42.
- Crisp, Jeff. 2021. "The State of the World's Refugees: Crisis or Progress?" *Journal of Human Security Studies* 10 (2): 1–14.
- Dafoe, Allan. 2014. "Science Deserves Better: The Imperative to Share Complete Replication Files." *PS: Political Science & Politics* 47 (1): 60–6.
- Davenport, Christian, Will Moore, and Steven Poe. 2003. "Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964–1989." *International Interactions* 29 (1): 27–55.
- Duvendack, Maren, Richard Palmer-Jones, and W. Robert Reed. 2017. "What Is Meant by 'Replication' and Why Does It Encounter Resistance in Economics?" *American Economic Review* 107 (5): 46–51.
- Echevarria, Jon, and Javier Gardeazabal. 2021. "A Spatial Model of Internal Displacement and Forced Migration." *Journal of Conflict Resolution* 25 (2–3): 591–618.
- Esarey, Justin, and Ahra Wu. 2016. "Measuring the Effects of Publication Bias in Political Science." *Research & Politics* 3 (3). <https://doi.org/10.1177/2053168016665856>.
- Fearon, James D., and Andrew Shaver. 2020. "Civil War Violence and Refugee Outflows." ESOC Working Paper Series (No. 25).
- Fisher, Jonathan D., David S. Johnson, and Timothy M. Smeeding. 2013. "Measuring the Trends in Inequality of Individuals and Families: Income and Consumption." *American Economic Review* 103 (3): 184–88.
- Freese, Jeremy, and David Peterson. 2017. "Replication in Social Science." *Annual Review of Sociology* 43: 147–65.
- Gerber, Alan, and Neil Malhotra. 2008. "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* 3 (3): 313–26.
- Hamermesh, Daniel S. 2007. "Replication in Economics." *Canadian Journal of Economics/Revue Canadienne d'économie* 40 (3): 715–33.
- King, Gary. 2006. "Publication, Publication." *PS: Political Science & Politics* 39 (1): 119–25.
- Marbach, Moritz. 2018. "On Imputing UNHCR Data." *Research & Politics* 5 (4). <https://doi.org/10.1177/2053168018803239>.
- Milton, Daniel, Megan Spencer, and Michael Findley. 2013. "Radicalism of the Hopeless: Refugee Flows and Transnational Terrorism." *International Interactions* 39 (5): 621–45.
- Moore, Will H., and Stephen M. Shellman. 2004. "Fear of Persecution: Forced Migration, 1952–1995." *Journal of Conflict Resolution* 48 (5): 723–45.
- Moore, Will H., and Stephen M. Shellman. 2007. "Whither Will They Go? A Global Study of Refugees' Destinations, 1965–1995." *International Studies Quarterly* 51 (4): 811–34.
- Moorthy, Shweta, and Robert Brathwaite. 2019. "Refugees and Rivals: The International Dynamics of Refugee Flows." *Conflict Management and Peace Science* 36 (2): 131–48.
- Nyrup, Jacob, and Stuart Bramwell. 2020. "Who Governs? A New Global Dataset on Members of Cabinets." *American Political Science Review* 114 (4): 1366–74.
- Pesaran, Hashem. 2003. "Introducing a Replication Section." *Journal of Applied Econometrics* 18 (1): 111.
- Polo, Sara M. T., and Julian Wucherpfennig. 2022. "Trojan Horse, Copycat, or Scapegoat? Unpacking the Refugees-Terrorism Nexus." *Journal of Politics* 84 (1): 33–49.
- Rubin, Jacqueline H., and Will H. Moore. 2007. "Risk Factors for Forced Migrant Flight." *Conflict Management and Peace Science* 24 (2): 85–104.
- Salehyan, Idean. 2008. "The Externalities of Civil Strife: Refugees as a Source of International Conflict." *American Journal of Political Science* 52 (4): 787–801.
- Salehyan, Idean, and Kristian Skrede Gleditsch. 2006. "Refugees and the Spread of Civil War." *International Organization* 60 (2): 335–66.
- Savun, Burcu, and Christian Gineste. 2019. "From Protection to Persecution: Threat Environment and Refugee Scapegoating." *Journal of Peace Research* 56 (1): 88–102.
- Shaver, Andrew, Benjamin Krick, Judy Blancaflor, Xavier Liu, Ghassan Samara, Sarah Yein Ku, Shengkuo Hu, et al. 2024. "Replication Data for: The Causes and Consequences of Refugee Flows: A Contemporary Reanalysis." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/JADOZL>.
- Simmons, Beth A., and Michael R. Kenwick. 2022. "Border Orientation in a Globalizing World." *American Journal of Political Science* 66 (4): 853–70.
- Steele, Abbey. 2017. *Democracy and Displacement in Colombia's Civil War*. Ithaca, NY: Cornell University Press.
- Turkdoglu, Oguzhan, and Thomas Chadeaux. 2019. "Nowhere to Go? Why Do Some Civil Wars Generate More Refugees Than Others?" *International Interactions* 45 (2): 401–20.
- UNHCR. 2021a. "UNHCR Refugee Population Statistics Database Solutions Dataset."
- UNHCR. 2021b. "UNHCR Refugee Population Statistics Database: Forced Displacement Flow Dataset."
- Wright, Thorin M., and Shweta Moorthy. 2018. "Refugees, Economic Capacity, and Host State Repression." *International Interactions* 44 (1): 132–55.