Forecasting International Migration to North Dakota: A Multi-Method Empirical Analysis

[Author Name][1]

[1][Institutional Affiliation]

[Email Address]

December 31, 2025

Abstract

International migration to small U.S. states presents distinctive forecasting challenges characterized by high volatility, limited sample sizes, and heightened sensitivity to federal policy changes. While the demographic literature has extensively examined migration to gateway states, small peripheral states—where migration often determines whether populations grow or decline—have received comparatively little rigorous empirical attention. This study addresses this gap through the first comprehensive multi-method analysis of international migration to North Dakota, deploying nine analytical modules spanning descriptive statistics, time series econometrics, panel regression, gravity models, machine learning, causal inference, and duration analysis using data from 2002–2024.

Key findings reveal that North Dakota's migration profile is dominated by refugee resettlement rather than economic migration chains, with Location Quotients exceeding 15 for certain African and Middle Eastern origins. The time series exhibits I(1) integration with a coefficient of variation of 82.5%, indicating that annual flows can deviate from expectations by amounts approaching the mean itself. Structural break analysis identifies significant regime shifts in 2020 and 2021 associated with the COVID-19 pandemic ($F = 16.01$, $p < 0.001$).

Difference-in-differences estimation identifies a policy-associated divergence of approximately 75% in refugee arrivals from Travel Ban-affected countries ($p = 0.032$), an effect that aggregate time series analysis fails to detect; however, pre-treatment trend divergence suggests this estimate likely represents an upper bound on the causal effect. Gravity model estimation reveals that diaspora associations lose statistical significance when controlling for population mass (coefficient attenuates from 0.45 to 0.14, $p = 0.57$), suggesting diaspora-linked allocation patterns remain modest in this cross-sectional specification.

Monte Carlo scenario projections through 2045 yield a median net international migration forecast of 9,056 persons with a 95% prediction interval spanning 3,570 to 14,491—a factor of 4.1—honestly characterizing the structural uncertainty inherent in small-state migration forecasting. This wide interval has direct implications for demographic planning: agencies must develop contingent capacity strategies rather than relying on misleadingly precise point forecasts. The multi-method framework demonstrates that rigorous demographic analysis remains feasible even when small populations strain conventional asymptotic methods, provided analysts attend carefully to identification assumptions and uncertainty quantification.

Keywords: international migration, forecasting, North Dakota, causal inference, demographic projection, refugee resettlement, Great Plains

JEL Codes: J11, J61, C22, C53, R23

1 Introduction

North Dakota occupies a distinctive position in the landscape of American immigration. With a population of approximately 780,000—ranking 47th among U.S. states—the state receives just 0.17% of national international migration flows, a proportion noticeably smaller than its 0.23% share of the U.S. population (Wilson et al., 2022). Yet this modest arithmetic masks a demographic reality of considerable consequence. In a state where natural population growth has stagnated and domestic out-migration persists, international migration has emerged as a critical—and at times dominant—component of population change. Understanding the patterns, drivers, and future trajectories of this migration stream carries implications not only for North Dakota's demographic planning but also for the broader challenge of sustaining communities in America's rural heartland.

The Great Plains has experienced sustained population decline throughout much of the twentieth and early twenty-first centuries, a demographic trajectory shaped by agricultural consolidation, limited economic diversification, and the magnetic pull of metropolitan opportunity elsewhere (Albrecht, 1993; Archer and Lonsdale, 2003). North Dakota's brief reversal of this trend during the Bakken oil boom (2010–2014) demonstrated how economic shocks can temporarily reshape demographic patterns, yet the underlying structural forces favoring out-migration have proven resilient (Weber et al., 2014). Against this backdrop, international migration—particularly refugee resettlement—has assumed an outsized role in maintaining population levels and meeting labor force demands in communities throughout the state.

Refugee resettlement programs have channeled arrivals from conflict zones worldwide to North Dakota communities, creating ethnic enclaves that now span generations. The state's foreign-born population reflects this history: Location Quotient analysis reveals that North Dakota hosts disproportionately large shares of migrants from East African origins (Somalia, Sudan, Ethiopia) and Middle Eastern countries (Iraq, Syria) relative to national patterns. These concentrations, products of federal resettlement policy interacting with local voluntary agency capacity, create diaspora networks that are predictive of subsequent migration streams (Beine et al., 2011; Bansak et al., 2018). The result is a migration profile markedly different from gateway states, one characterized by humanitarian flows rather than economic migration chains connecting to traditional sending regions.

Despite the demographic significance of international migration to small states and rural regions, the empirical literature on state-level migration forecasting remains surprisingly thin. The dominant paradigm in migration forecasting focuses on national-level flows, employing gravity models and time series methods calibrated to large-sample settings (Massey et al., 1993; Mayda, 2010). When subnational analysis appears, it typically examines major receiving states—California, Texas, Florida, New York—where sample sizes support conventional inferential approaches (Portes and Rumbaut, 2006). Small states, characterized by volatile flows, limited historical observations, and heightened sensitivity to policy shocks, have largely escaped systematic empirical scrutiny.

This lacuna reflects genuine methodological challenges. International migration to North Dakota

exhibits a coefficient of variation of 82.5%, reflecting annual flows that have ranged from 30 to over 5,000 persons within a fifteen-year span. Time series of this brevity strain the asymptotic foundations of standard econometric procedures; unit root tests lose power, ARIMA model selection becomes unreliable, and long-horizon forecasts carry prediction intervals so wide as to approach uninformativeness (Cerqueira et al., 2019; Hyndman and Athanasopoulos, 2021). The challenge, then, is not merely to apply established methods but to develop an analytical framework suited to the small-sample, high-volatility setting that characterizes migration to peripheral regions. Such a framework must integrate multiple methodological traditions, triangulating findings across approaches while honestly characterizing the uncertainty inherent in projection.

This analysis addresses four interconnected research questions regarding international migration to North Dakota:

1. **Patterns and Sources**: What are the dominant patterns and geographic sources of international migration to North Dakota, and how has the composition of this migration stream evolved over time?

2. **Time Series Properties**: What statistical properties characterize the international migration time series—specifically, what is the order of integration, are structural breaks present, and what degree of persistence or mean reversion do the data exhibit?

3. **Policy Effects**: What are the effects of major policy interventions—the 2017 Travel Ban and the 2020 COVID-19 pandemic response—on international migration flows to North Dakota, and what are the bounds on causal interpretation?

4. **Future Scenarios**: Given observed patterns and structural uncertainty, what is the range of plausible future migration scenarios through 2045, and how should demographic planners characterize forecast uncertainty?

These questions progress from descriptive characterization through inferential analysis to applied forecasting, reflecting the logical structure of demographic inquiry. Each question demands distinct methodological approaches, and the answers collectively inform the population projection enterprise.

This study makes three principal contributions to the migration forecasting literature. First, it provides the first comprehensive multi-method empirical analysis of international migration to a small U.S. state. The analysis deploys nine interconnected methodological modules spanning descriptive statistics, time series analysis, panel regression, gravity models, machine learning, causal inference, and duration analysis. This methodological breadth serves not eclecticism but triangulation: findings that emerge consistently across disparate analytical frameworks warrant greater confidence than those dependent on particular modeling assumptions. The multi-method design explicitly addresses the small-sample challenge by trading depth within any single paradigm for robustness across paradigms.

Second, the analysis offers novel applications of causal inference methods to state-level migration policy evaluation. Difference-in-differences estimation, applied to refugee arrivals before and after the 2017 Travel Ban, identifies a policy-associated divergence of approximately 75% in the first full post years ($p = 0.032$), an effect that aggregate time series analysis fails to detect. Because the joint pre-treatment test rejects parallel trends over the full pre-period, this estimate likely conflates the policy effect with pre-existing divergence and should be interpreted as an upper bound. Because the Travel Ban is a national shock, a synthetic comparator is used only as a descriptive benchmark for North Dakota's international migration rate, not a causal counterfactual. A shift-share index based on leave-one-out national shocks provides first-stage relevance linking refugee-driven inflows to state migration. These applications demonstrate that policy-sensitive inference remains feasible in small-state settings when identification assumptions are stated and bounded.

Third, the study develops a rigorous framework for uncertainty quantification in long-range migration forecasts. Monte Carlo simulation propagates parameter uncertainty through projection models, generating probability distributions rather than point estimates for future migration. The resulting 95% prediction interval for 2045 international migration spans 3,570 to 14,491 persons—a factor of 4.1 times—honestly reflecting the structural uncertainty that characterizes this domain. This probabilistic framing enables demographic planners to assess the full range of contingencies rather than anchoring on misleadingly precise point forecasts (Raftery et al., 2014). The scenario analysis further distinguishes between forecast uncertainty (arising from model imprecision) and scenario uncertainty (arising from unknowable future policy regimes), providing a framework for structured demographic planning under deep uncertainty.

The remainder of this article proceeds as follows. Section 2 describes the data sources—Census Bureau Population Estimates, Department of Homeland Security admission records, American Community Survey foreign-born tabulations, and refugee arrival databases—and presents the methodological framework in detail. Section 3 reports findings organized by analytical module, beginning with descriptive patterns and concentration analysis, proceeding through time series diagnostics and panel regression results, and culminating in causal estimates and scenario projections. Section 4 interprets these findings in light of migration theory and prior empirical work, addresses policy implications for refugee resettlement and demographic planning, and acknowledges the limitations inherent in small-sample analysis. Section 5 summarizes the contributions and identifies directions for future research.

2 Data and Methods

This section describes the data sources underlying the analysis and presents the methodological framework employed across nine analytical modules. The multi-method design reflects both the complexity of international migration dynamics and the inferential challenges posed by small-sample settings. Each methodological tradition contributes distinct insights: descriptive methods establish baseline patterns; time series techniques characterize dynamic properties; regression approaches

quantify relationships; and causal inference methods identify policy effects. Together, these approaches enable triangulation across analytical frameworks, enhancing confidence in findings that emerge consistently while flagging conclusions dependent on particular modeling assumptions.

2.1 Analytical Pipeline Overview

Figure 1 summarizes the nine-module pipeline and the flow of information into the scenario engine. Modules 1–6 provide descriptive diagnostics, baseline time-series structure, and predictive covariates. Module 7 offers causal evidence of policy sensitivity, which informs the narrative interpretation of scenarios but does not replace the forecast target. Module 8 translates refugee wave persistence into a forecasting adjustment, and Module 9 integrates all inputs into scenario paths and Monte Carlo uncertainty.

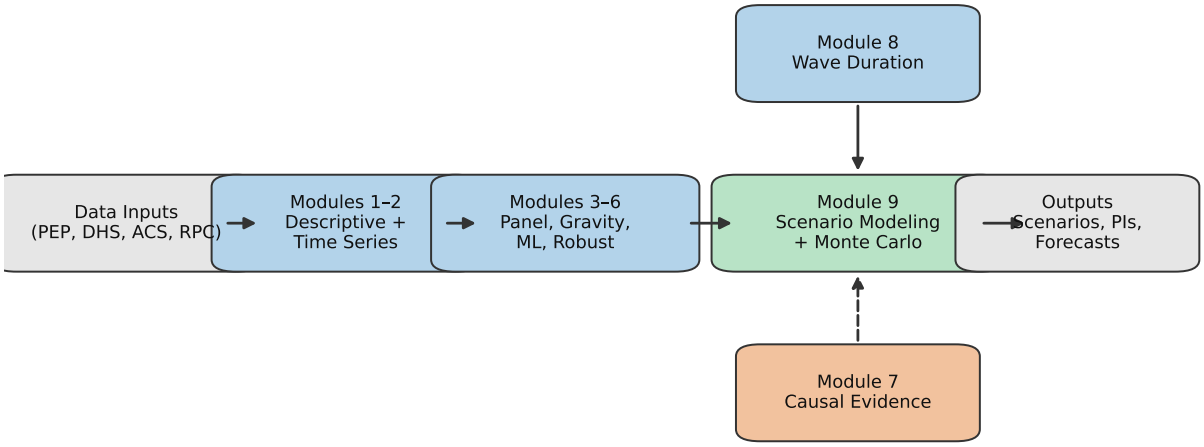

Figure 1: Analytical pipeline linking data sources, module groups, and scenario outputs. Solid arrows denote the forecasting flow; dashed arrows indicate policy-evidence inputs.

2.2 Data Sources

The analysis draws on four primary data sources, each capturing distinct dimensions of international migration to North Dakota. We define a single forecasting estimand based on the Census Bureau Population Estimates Program (PEP) net international migration component for North Dakota, and treat all other sources as auxiliary inputs for decomposition, prediction, and validation. Table 2 maps each source to its measurement properties (flow versus stock; net versus gross; time basis) and its role relative to the forecast target.

Table 1: Module classification and role in the forecasting pipeline

Modules	Primary focus	Role
1.1–1.2	Descriptive statistics and concentration	Forecasting context
2.1–2.2	Time-series diagnostics and baselines	Forecasting context
3.1–4	Panel and robust regression	Forecasting inputs
5	Gravity and diaspora associations	Forecasting inputs
6	Machine learning predictors	Forecasting inputs
7	Policy evaluation (DiD, ITS)	Causal evidence
8	Wave duration modeling	Forecasting input
9	Scenario modeling and Monte Carlo	Forecasting output

Notes: Forecasting-role modules feed scenario construction; causal evidence informs policy sensitivity but does not redefine the estimand.

2.2.1 Estimand & Measurement

Forecast target (estimand). Let Y_t denote North Dakota's annual *net international migration* in year t, measured as the PEP "international migration" component of change for North Dakota (persons; net).[1] We treat Y_t as the primary dependent variable for forecasting and scenario projections in this paper. Unless otherwise noted, the generic notation y_t used in subsequent module equations denotes the observed PEPnet international migration series. Our forecasting goal is to characterize the predictive distribution of Y_{t+h} for horizons $h \in \{1, \ldots, H\}$ (2025–2045), conditional on information available through year t.

Auxiliary data sources. DHS lawful permanent resident (LPR) admissions, Refugee Processing Center (RPC) refugee arrivals, and ACSforeign-born stock estimates are not alternative outcomes; they are used to (i) characterize and partially decompose international migration, (ii) construct predictors (e.g., diaspora measures and origin-composition features), and (iii) triangulate mechanisms and scenario sensitivity.

Time base and harmonization. PEPestimates are reported on an estimate-year basis, while DHS and RPC series are reported in fiscal years and ACSestimates in survey years. We preserve native time bases and label them explicitly; when fiscal-year series are used as predictors for PEP-year outcomes, we align them using an overlap-weighted crosswalk $X_t^{\text{PEP}} \approx 0.75\, X_t^{\text{FY}} + 0.25\, X_{t-1}^{\text{FY}}$, and we report sensitivity to one-year shifts where timing is ambiguous.

These sources are related but not interchangeable: PEPinternational migration is a *net* component of population accounting, DHS and RPC series are *gross inflows* on a fiscal-year basis, and ACSseries are *stocks* subject to sampling error. Throughout the paper we label flow/stock status and time base (FY vs. PEP-year vs. survey year) in text, tables, and figure captions to prevent target drift.

[1] PEP "year t" corresponds to the annual component-of-change interval used in producing July 1 population estimates (a demographic year rather than a strict January–December calendar year). We refer to it as "year t" for readability and explicitly label fiscal-year sources as FY.

Table 2: Data Source Mapping Relative to the Forecast Target Y_t

Source	**What it measures**	**Flow Stock**	**Net Gross**	**Time basis (coverage)**	**Role relative to estimand and key caveats**
Census PEP (components of change)	State net international migration component of population change (persons)	Flow	Net	PEP estimate year (2010–2024)	**Primary forecast target Y_t.** Model-based estimates; composite of multiple international channels; subject to revision; 2020 reflects both true disruption and potential measurement artifacts.
RPC refugee arrivals	Refugee (and related humanitarian) arrivals by nationality and state of initial placement	Flow	Gross	Fiscal year (FY; 2002–2020)	Mechanism and scenario input for the refugee component; supports policy responsiveness analysis. Initial placement $\neq$ final residence (secondary migration); FY timing differs from PEP; many zeros at state×nationality.
DHS LPRadmissions	Lawful permanent resident admissions by origin (state of intended residence)	Flow	Gross	Fiscal year (FY; 2023)	Predictor and composition diagnostics (origin structure; gravity-style features). Intended residence may differ from actual; single cross-section; excludes other statuses; not net of outflows.
ACS foreign-born stock	Foreign-born population stock by place of birth (survey estimate; MOE)	Stock	—	Survey year (5-year est.; 2009–2023)	Diaspora predictors and validation for origin concentration. Sampling error (MOE) is nontrivial for small cells; stock conflates past flows, retention, and internal migration; measurement error can attenuate relationships.

2.2.2 Census Bureau Population Estimates Program

The U.S. Census Bureau's Population Estimates Program (PEP) provides annual estimates of population and components of change for states and counties. This analysis employs the vintage 2024 estimates covering 2010–2024, which decompose annual population change into births, deaths, domestic migration, and international migration. The international migration component—designated `INTERNATIONALMIG` in Census Bureau nomenclature—comprises the net flow of foreign-born individuals crossing U.S. borders plus the net movement of native-born citizens and military personnel abroad. For subnational units, the Census Bureau allocates national international migration estimates to states using American Community Survey data on recent movers and foreign-born population distributions.

The PEP data provide the primary dependent variable for time series analysis: annual net international migration to North Dakota. Key advantages include consistent methodology across years and comprehensive geographic coverage. Limitations include the composite nature of the international migration variable, which aggregates heterogeneous streams (lawful permanent residents, refugees, temporary workers, and unauthorized migrants) into a single measure, and the allocation methodology, which may introduce smoothing artifacts at the state level.

2.2.3 Department of Homeland Security Lawful Permanent Resident Data

The Department of Homeland Security (DHS) publishes annual statistics on lawful permanent resident (LPR) admissions by state of intended residence and country of birth. This analysis employs fiscal year 2023 data, the most recent available at the time of analysis. The LPR data provide granular information on the country-of-origin composition of legal immigration flows, enabling gravity model estimation and concentration analysis.

The principal advantage of DHS data lies in their administrative provenance: unlike survey-based measures, LPR counts derive from visa issuance records and represent actual admissions rather than estimates. Limitations include restriction to a single cross-section (precluding panel analysis within this data source), exclusion of non-LPR categories (refugees, asylees, temporary workers), and the distinction between state of intended residence at admission and actual destination after secondary migration.

2.2.4 American Community Survey Foreign-Born Population

The American Community Survey (ACS) provides annual estimates of foreign-born population by place of birth for states and metropolitan areas. This analysis employs five-year estimates from 2009–2023, using the detailed country-of-birth tabulations (Table B05006). The foreign-born stock data enable analysis of diaspora effects on new migration, origin-specific growth trajectories, and Location Quotient calculations comparing North Dakota's immigrant composition to national patterns.

ACS data carry sampling uncertainty quantified through published margins of error. For small

populations—such as foreign-born from specific countries residing in North Dakota—coefficients of variation can exceed 30%, requiring cautious interpretation. The analysis addresses this limitation by focusing on origin groups with sufficient sample sizes and by aggregating to regional categories where country-level estimates prove unreliable.

2.2.5 Refugee Processing Center Arrival Data

The Refugee Processing Center (RPC), operated by the Department of State, maintains comprehensive records of refugee arrivals by state of initial resettlement and nationality. This analysis employs arrival data spanning fiscal years 2002–2020, providing the longest time series among the data sources and enabling duration analysis of migration "waves" from specific origin countries. The refugee data prove particularly valuable for North Dakota given the state's significant refugee population, which drives much of the geographic concentration observed in other data sources.

Refugee arrival data offer precise counts (not estimates) of a well-defined population admitted through federal resettlement programs. Limitations include truncation at fiscal year 2020 (precluding analysis of pandemic recovery) and restriction to initial placement, which may differ from eventual settlement after secondary migration.

2.3 Descriptive and Concentration Methods

Descriptive analysis establishes baseline patterns in international migration flows and characterizes their distributional properties. Summary statistics—means, standard deviations, coefficients of variation, skewness, and kurtosis—quantify central tendency and dispersion. The coefficient of variation (CV), defined as the ratio of standard deviation to mean, provides a scale-invariant measure of volatility suitable for comparing series of different magnitudes.

To decompose observed migration flows into trend and cyclical components, the analysis employs the Hodrick-Prescott (HP) filter (Hodrick and Prescott, 1997). The HP filter solves the optimization problem:

$$\min_{\{g_t\}} \left\{ \sum_{t=1}^{T} (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right\} \tag{1}$$

where y_t denotes the observed series, g_t the trend component, and λ the smoothing parameter controlling the penalty on trend acceleration. Following Ravn and Uhlig (2002), the analysis sets $\lambda = 6.25$ for annual data, a value calibrated to achieve comparable smoothing across observation frequencies.

Geographic concentration analysis employs two complementary measures. The Herfindahl-Hirschman Index (HHI) quantifies concentration across origin countries:

$$\text{HHI} = \sum_{i=1}^{N} s_i^2 \times 10{,}000 \tag{2}$$

where s_i denotes the share of migration from origin country i. HHI values below 1,500 indicate

unconcentrated distributions; values between 1,500 and 2,500 indicate moderate concentration; values above 2,500 indicate high concentration. Location Quotients (LQ) measure the degree to which North Dakota's immigrant composition deviates from national patterns:

$$\text{LQ}_{i,\text{ND}} = \frac{(\text{Foreign-born from } i \text{ in ND})/(\text{Total foreign-born in ND})}{(\text{Foreign-born from } i \text{ in US})/(\text{Total foreign-born in US})} \tag{3}$$

Location quotients exceeding unity indicate overrepresentation of origin group i in North Dakota relative to the nation.

2.4 Time Series Methods

Time series analysis characterizes the stochastic properties of international migration flows and provides a foundation for forecasting. The analysis proceeds through four stages: unit root testing, ARIMA model specification, structural break detection, and vector autoregression for multivariate dynamics.

2.4.1 Unit Root Tests

Stationarity assessment employs ADF and PP tests alongside KPSS (null: stationarity) to triangulate integration properties in a very short series. The Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) estimates the regression:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + \varepsilon_t \tag{4}$$

and tests $H_0 : \gamma = 0$ (unit root) against $H_1 : \gamma < 0$ (stationarity). Lag length p is selected by the Akaike Information Criterion (AIC). The Phillips-Perron (PP) test (Phillips and Perron, 1988) provides a nonparametric alternative that corrects for serial correlation without augmenting the regression. KPSS tests are run with a constant-only specification (level-stationarity). Break-robust Zivot-Andrews tests (single endogenous break) are reported as sensitivity checks. Where tests yield conflicting conclusions, the analysis reports both and interprets results cautiously.

2.4.2 ARIMA Model Selection

For series exhibiting unit roots, the analysis fits autoregressive integrated moving average (ARIMA) models of order (p, d, q), where d denotes the differencing order required for stationarity (Box and Jenkins, 1970). Model selection employs the AIC criterion:

$$\text{AIC} = -2\ln(\hat{L}) + 2k \tag{5}$$

where $\hat{L}$ denotes the maximized likelihood and k the number of parameters. Diagnostic checking employs the Ljung-Box portmanteau test for residual autocorrelation and the Shapiro-Wilk test for

residual normality.

2.4.3 Structural Break Tests

Structural break analysis employs three complementary approaches. The Chow test (Chow, 1960) evaluates whether regression parameters differ across known candidate break points by comparing residual sums of squares from restricted (single-regime) and unrestricted (regime-specific) models. The CUSUM test (Brown et al., 1975) detects parameter instability through cumulative sums of recursive residuals; departures from the zero line indicate structural change. For endogenous break detection, the analysis employs the Bai and Perron (1998) procedure, which identifies multiple break points by minimizing the sum of squared residuals across all possible segmentations:

$$(\hat{T}_1, \ldots, \hat{T}_m) = \arg\min_{T_1,\ldots,T_m} \sum_{j=0}^{m} \sum_{t=T_j+1}^{T_{j+1}} (y_t - \bar{y}_j)^2 \tag{6}$$

subject to minimum segment length constraints. Given the short time series available (n = 15), the analysis sets the minimum segment length to 2 years, acknowledging the resulting power limitations.

2.4.4 Vector Autoregression

To examine dynamic interdependencies between North Dakota and national migration flows, the analysis estimates a vector autoregression (VAR) of order p:

$$\mathbf{y}_t = \mathbf{c} + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t \tag{7}$$

where $\mathbf{y}_t = (y_t^{\text{ND}}, y_t^{\text{US}})'$ contains North Dakota and national international migration, $\mathbf{A}_i$ are coefficient matrices, and $\boldsymbol{\varepsilon}_t$ is a vector white noise process. Lag order selection employs AIC, and Granger causality tests evaluate directional predictive relationships. Forecast error variance decomposition quantifies the contribution of each variable to forecast uncertainty at various horizons.

2.5 Panel Data Methods

Panel data analysis exploits variation across states and time to estimate relationships unidentifiable from single-state time series. The analysis employs state-year observations spanning 51 states (including the District of Columbia) over 15 years (2010–2024), yielding 765 observations.

The fixed effects model specifies:

$$y_{it} = \alpha_i + \lambda_t + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \tag{8}$$

where y_{it} denotes international migration to state i in year t, α_i represents state-specific intercepts absorbing time-invariant heterogeneity, λ_t captures common year effects, and $\mathbf{x}_{it}$ contains

time-varying covariates. The random effects model treats α_i as random draws from a population distribution rather than fixed parameters to be estimated.

Model selection between fixed and random effects employs the Hausman test (Hausman, 1978), which evaluates whether random effects estimates exhibit bias due to correlation between α_i and regressors. Under the null hypothesis of no correlation, random effects is efficient; under the alternative, fixed effects is consistent but random effects is biased. The test statistic follows a chi-squared distribution under the null:

$$H = (\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{RE}})'[\text{Var}(\hat{\boldsymbol{\beta}}_{\text{FE}}) - \text{Var}(\hat{\boldsymbol{\beta}}_{\text{RE}})]^{-1}(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{RE}}) \sim \chi^2_k \tag{9}$$

Standard errors are clustered at the state level to account for within-state serial correlation.

2.6 Cross-Sectional Allocation Models

Cross-sectional allocation models quantify how LPR admissions are distributed across state–origin pairs as a function of diaspora stocks and destination characteristics. Although inspired by the gravity framework (Tinbergen, 1962; Anderson and van Wincoop, 2003), our specification differs in two respects: (1) it uses a single FY2023 cross-section rather than bilateral panel data, and (2) it omits origin–destination distance because our destination units are U.S. states receiving flows from the same origin countries—variation in distance to a given origin is minimal and dominated by within-U.S. allocation patterns rather than international migration costs.[2]

The analysis estimates specifications using the Poisson pseudo-maximum likelihood (PPML) estimator advocated by Santos Silva and Tenreyro (2006), which addresses two limitations of log-linear OLS: heteroskedasticity-induced bias and the inability to accommodate zero flows.

The estimated specification takes the form:

$$E[M_{od}] = \exp(\beta_0 + \beta_1 \ln \text{Stock}_{od} + \beta_2 \ln \text{OriginTotal}_o + \beta_3 \ln \text{DestTotal}_d + \boldsymbol{\gamma}'\mathbf{Z}_{od}) \tag{10}$$

where M_{od} denotes LPR admissions from origin country o to destination state d (FY2023), Stock_{od} is the existing diaspora (foreign-born from o residing in d), OriginTotal_o is the national foreign-born population from origin o, DestTotal_d is the total foreign-born population in destination state d, and $\mathbf{Z}_{od}$ contains additional bilateral controls. The coefficient β_1 on diaspora stock captures a cross-sectional diaspora–flow association useful for prediction.

Diaspora elasticity estimation acknowledges endogeneity concerns: diaspora stocks and migration flows are jointly determined, and historical flows mechanically generate stocks. The analysis therefore frames coefficients as predictive associations rather than causal effects; robustness checks consider lagged stock measures and historical settlement patterns (Card, 2001).

[2]Traditional gravity models estimate distance effects using cross-country bilateral flows where distance proxies transaction costs. In our state-level allocation context, the relevant "distance" would be origin–state distance, which is nearly constant across states for a given origin (e.g., distance from Nepal to North Dakota versus to Minnesota differs by less than 2%). We therefore focus on diaspora networks and destination scale as the primary allocation mechanisms.

2.7 Machine Learning Methods

Machine learning methods complement traditional regression by providing flexible functional forms and systematic variable selection. The analysis employs three techniques: Elastic Net regularization, Random Forest regression, and K-means clustering.

Elastic Net (Zou and Hastie, 2005) combines L_1 (Lasso) and L_2 (Ridge) penalties, solving:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} (y_i - \mathbf{x}_i'\boldsymbol{\beta})^2 + \lambda \left[\alpha \|\boldsymbol{\beta}\|_1 + (1-\alpha)\|\boldsymbol{\beta}\|_2^2 \right] \right\} \tag{11}$$

where λ controls overall regularization strength and $\alpha \in [0, 1]$ balances sparsity-inducing L_1 penalty against coefficient-shrinking L_2 penalty. Cross-validation selects (λ, α) values. Coefficients shrunk exactly to zero identify variables excluded from the predictive model.

Random Forest (Breiman, 2001) aggregates predictions from an ensemble of decision trees, each trained on bootstrap samples with random feature subsets. Feature importance is measured by permutation importance: the decrease in out-of-bag prediction accuracy when feature values are randomly shuffled.

K-means clustering partitions states into groups exhibiting similar migration profiles. The algorithm minimizes within-cluster sum of squared distances to cluster centroids:

$$\arg\min_{\mathcal{C}} \sum_{k=1}^{K} \sum_{i \in C_k} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|^2 \tag{12}$$

where C_k denotes cluster k and $\boldsymbol{\mu}_k$ its centroid. Optimal cluster count K is selected by the silhouette criterion.

2.8 Causal Inference Methods

Causal inference methods estimate the effects of specific policy interventions—the 2017 Travel Ban and the 2020 COVID-19 pandemic—on international migration flows. Identification relies on nationality-level difference-in-differences and a state-level interrupted time series. Two additional components provide context: a synthetic comparator for descriptive benchmarking and a shift-share index used to assess the predictive relevance of refugee-driven shocks.

2.8.1 Difference-in-Differences

The Travel Ban analysis employs a difference-in-differences (DiD) design comparing refugee arrivals from affected countries (Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen) to arrivals from unaffected countries, before and after the 2017 implementation. The estimating equation specifies:

$$\ln(y_{ct} + 1) = \alpha_c + \lambda_t + \delta \cdot (\text{Affected}_c \times \text{Post}_t) + \varepsilon_{ct} \tag{13}$$

where y_{ct} denotes arrivals from country c in year t, α_c and λ_t are country and year fixed effects, and the coefficient δ identifies the average treatment effect on the treated (ATT). The logarithmic transformation accommodates the multiplicative nature of policy effects; the addition of unity handles zero arrivals.

Identification requires the parallel trends assumption: absent the Travel Ban, affected and unaffected countries would have exhibited similar trends in arrivals. The analysis evaluates this assumption through pre-treatment trend tests and graphical inspection of pre-period trajectories. Standard errors are clustered by nationality to account for serial correlation within origin groups.

Small-sample inference robustness. With only seven treated nationalities, conventional cluster-robust inference may be unreliable (Cameron et al., 2008). We supplement clustered standard errors with two small-sample approaches. First, wild cluster bootstrap with Rademacher weights generates the null distribution by resampling cluster-level residuals under the null hypothesis of no treatment effect; bootstrap p-values correct for the small number of treated clusters. Second, randomization (permutation) inference implements Fisher's exact test by permuting the treatment assignment across clusters and computing the proportion of permutations yielding an ATT as extreme as observed. Both approaches provide finite-sample valid inference without relying on asymptotic cluster count assumptions.

Restricted pre-period robustness. When the joint pre-trend test rejects parallel trends over the full pre-period, we re-estimate the DiD restricting to years where pre-trends appear more plausibly parallel. If event study coefficients show greater divergence in early years than in years immediately preceding treatment, restricting the pre-period strengthens the parallel trends assumption at the cost of reduced sample size.

2.8.2 Interrupted Time Series (COVID-19)

The COVID-19 analysis uses a state-level interrupted time series with state fixed effects to identify both an immediate level shift and a post-2020 trend change:

$$y_{st} = \alpha_s + \beta_1 t + \beta_2 \text{Post}_{2020,t} + \beta_3 (t - 2020)\text{Post}_{2020,t} + \varepsilon_{st} \tag{14}$$

where y_{st} is net international migration for state s in year t. The coefficient β_2 captures the immediate COVID-19 level shift, and β_3 captures the change in trend after 2020. Standard errors are clustered by state.

2.8.3 Synthetic Comparator (Descriptive Benchmark)

Because the Travel Ban is a national shock, donor states are not untreated. We therefore use a synthetic comparator (Abadie and Gardeazabal, 2003; Abadie et al., 2010) only as a descriptive benchmark. The comparator trajectory is constructed by weighting donor states to match North

Dakota's pre-2017 international migration rate:

$$\hat{y}_{1t}^{S} = \sum_{j=2}^{J+1} w_j^* y_{jt} \tag{15}$$

with weights $w_j^* \geq 0$, $\sum_j w_j^* = 1$ chosen to minimize pre-treatment discrepancy. Post-2017 gaps $y_{1t} - \hat{y}_{1t}^{S}$ are reported descriptively rather than interpreted as causal effects.

2.8.4 Shift-Share (Bartik) Index

To assess the predictive relevance of refugee-driven shocks for state-level migration, the analysis constructs a shift-share (Bartik) index (Bartik, 1991; Goldsmith-Pinkham et al., 2020). The index combines national-level "shifts" in immigration by origin country with state-level "shares" reflecting historical settlement patterns:

$$B_{dt} = \sum_{o} \omega_{od,t_0} \cdot g_{o,t}^{\mathrm{US},-d} \tag{16}$$

where ω_{od,t_0} is origin o's share of state d's baseline arrivals and $g_{o,t}^{\mathrm{US},-d}$ is the leave-one-out national change in arrivals from origin o (excluding state d). Because no second-stage causal model is estimated, the results are interpreted as first-stage relevance: the strength of the relationship between predicted refugee inflows and state-level international migration. Standard errors are clustered by state; first-stage F-statistics assess instrument strength, with values exceeding 10 indicating adequate relevance.

2.9 Duration Analysis

Duration analysis examines the lifecycle of migration "waves"—sustained periods of elevated arrivals from specific origin countries. A wave is defined operationally as a period during which arrivals from a nationality exceed 150% of that nationality's baseline for two or more consecutive years. This definition balances sensitivity (detecting genuine surges) against specificity (avoiding false positives from random fluctuation).

The Kaplan-Meier estimator (Kaplan and Meier, 1958) provides nonparametric estimates of the survival function $S(t) = P(T > t)$, where T denotes wave duration:

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{17}$$

where d_i is the number of waves ending at time t_i and n_i is the number at risk. The log-rank test evaluates whether survival curves differ across strata (e.g., intensity quartiles, origin regions).

The Cox proportional hazards model (Cox, 1972) relates the hazard rate—the instantaneous probability of wave termination conditional on survival—to covariates:

$$h(t|\mathbf{x}) = h_0(t) \exp(\boldsymbol{\beta}'\mathbf{x}) \tag{18}$$

where $h_0(t)$ is the baseline hazard and $\exp(\beta_j)$ gives the hazard ratio associated with a one-unit increase in covariate x_j. Hazard ratios below unity indicate factors prolonging waves; ratios above unity indicate factors accelerating termination. The proportional hazards assumption—that hazard ratios remain constant over time—is evaluated through Schoenfeld residual tests. Model discrimination is assessed by the concordance index (C-statistic), which measures the probability that predicted hazards correctly rank pairs of observations by survival time.

2.10 Module-to-Scenario Integration

Table 3 summarizes how outputs from each analytical module feed into the scenario engine. This mapping clarifies the role of each module in the forecasting pipeline: some modules provide direct forecasting inputs (parameter estimates, uncertainty bounds), while others provide contextual evidence that informs scenario interpretation but does not alter the forecast mechanics.

As Table 3 indicates, the VAR model (Module 2.2) and machine learning methods (Module 6) serve primarily as auxiliary analyses. The VAR model contributes to AIC-based model averaging weights but is not used for standalone long-horizon forecasting due to the infeasibility of obtaining future national migration values. Similarly, ML methods (Elastic Net, Random Forest, K-means clustering) provide feature importance rankings and state clustering for comparative context, but do not directly parameterize the scenario trajectories. These modules support triangulation and validation rather than primary forecast generation.

2.11 Scenario Construction

Scenario analysis translates analytical findings into forward-looking projections while explicitly characterizing uncertainty. The analysis develops four policy-indexed scenarios spanning the range of plausible future trajectories:

1. **CBO Full**: Uses ARIMA point forecasts scaled by 1.1 for 2025–2029, then compounds at 8% annually from 2030 onward. This scenario represents an upper bound predicated on policy liberalization.

2. **Moderate**: Uses ARIMA point forecasts for 2025–2029, then applies a dampened trend (50% of the averaged robust trend estimate) through 2045. This scenario represents the central tendency forecast.

3. **Pre-2020 Trend**: Anchors at the 2019 level and extrapolates the 2010–2019 slope, treating COVID as a temporary deviation. This scenario provides a lower-bound estimate under restrictive policy assumptions.

4. **Zero**: Assumes complete cessation of international migration, providing a floor for population projection under extreme policy scenarios.

Table 3: Module Outputs and Scenario Integration

Module	**Key Output(s)**	**Role in Scenario Engine**	**Scenarios Affected**
1.1–1.2 (Descriptive)	Summary statistics; HHI; Location Quotients	Context for origin composition assumptions; validates concentration patterns	All (interpretation)
2.1 (ARIMA)	Baseline drift ($\hat{\beta}$); residual variance ($\hat{\sigma}^2$)	ARIMA point forecasts anchor Moderate and CBO scenarios; variance feeds MC simulation	Moderate, CBO Full
2.2 (VAR)	AIC model weights; FEVD proportions	Model averaging weights for ensemble; not used for standalone forecast	Model ensemble
3–4 (Panel/Robust)	State/year fixed effects; covariate elasticities	Validates ND position in state distribution; informs plausibility bounds	All (validation)
5 (Cross-sectional)	Diaspora elasticity (predictive)	Context for composition stability; not directly in forecast	None (context)
6 (ML)	Feature importance rankings; cluster membership	Auxiliary validation; identifies ND peer states for comparison	None (context)
7 (Policy)	DiD ATT bounds; ITS level/trend shifts	Informs restrictive vs. permissive scenario multipliers; policy sensitivity	Pre-2020 Trend
8 (Duration)	Cox survival probabilities; wave lifecycle shape	Wave persistence draws for MC; conditional survival by intensity/region	MC simulation
9 (Scenarios)	Combined trajectory paths; MC prediction intervals	Final output	All

Notes: "Context" modules provide interpretive background but do not mechanically alter scenario trajectories. "Forecasting input" modules contribute parameter estimates or distributional draws to the scenario engine. MC = Monte Carlo.

Monte Carlo simulation propagates uncertainty through projection models by sampling from estimated parameter distributions. The procedure draws 1,000 realizations of model parameters, generates projected trajectories for each draw, and summarizes the resulting distribution through percentiles. The 50% prediction interval spans the 25th to 75th percentiles; the 95% prediction interval spans the 2.5th to 97.5th percentiles. This approach quantifies forecast uncertainty arising from parameter estimation imprecision while holding the structural model fixed.

Model averaging combines forecasts from multiple specifications using AIC-derived weights:

$$w_m = \frac{\exp(-\Delta_m/2)}{\sum_{j=1}^{M} \exp(-\Delta_j/2)} \tag{19}$$

where $\Delta_m = \text{AIC}_m - \text{AIC}_{\text{min}}$ is the AIC difference between model m and the best-fitting model. Weighted averaging reduces dependence on any single specification while preserving interpretability (Hyndman and Athanasopoulos, 2021).

2.12 Implementation

All analyses were conducted in Python 3.11. Time series analysis employed `statsmodels` for ARIMA, VAR, and unit root tests. Panel data models were estimated using `linearmodels`. Machine learning methods used `scikit-learn` for Elastic Net, Random Forest, and clustering. Survival analysis employed the `lifelines` package for Kaplan-Meier estimation and Cox regression. Structural break detection used the `ruptures` package. Replication code and data are available in the supplementary materials.

3 Results

This section presents findings from the nine-module analytical pipeline, organized to parallel the methodological sequence described in Section 2. Results progress from descriptive characterization through time series analysis, regression-based modeling, causal inference, duration analysis, and scenario projections. Throughout, statistical significance is reported at conventional levels ($\alpha = 0.05$), with exact p-values provided where informative.

3.1 Descriptive Patterns

Table 4 presents summary statistics for PEPnet international migration (the PEPinternational migration component of change) to North Dakota and the United States over the 2010–2024 observation period. North Dakota's net international migration averaged 1,796 persons annually (SE = 383), with substantial volatility: the coefficient of variation (CV) reached 82.5%, indicating that the standard deviation approaches the mean in magnitude. By comparison, national net international migration exhibited somewhat lower volatility (CV = 73.5%), suggesting that state-level flows experience amplified fluctuation relative to aggregate national trends.

Table 4: Summary Statistics: Net International Migration (PEP), 2010–2024

Variable	Mean	SD	CV	Min	Max
ND Net International Migration (PEP)	1,796	1,482	82.5%	30	5,126
US Net International Migration (PEP)	1,010,744	743,049	73.5%	19,885	2,786,119
ND Share of US Net Int'l Mig. (%)	0.173	0.054	31.3%	0.102	0.303
ND Share of US Pop. (%)	0.231	0.006	2.4%	0.218	0.236

Notes: $n = 15$ annual observations. SD = standard deviation; CV = coefficient of variation (SD/Mean). Net international migration data from Census Bureau PEP(international migration component of change) vintage 2024.

The distribution of annual international migration to North Dakota exhibits positive skewness (1.10), reflecting a right tail associated with recent high-migration years. The Shapiro-Wilk test marginally fails to reject normality ($W = 0.886$, $p = 0.058$), though this finding warrants cautious interpretation given the small sample size. No observations qualify as extreme outliers under the interquartile range criterion, though the 2020 value of 30 migrants represents a dramatic departure from typical levels.

Hodrick-Prescott filter decomposition reveals a U-shaped trend in North Dakota's share of national international migration. The trend component declined from 0.211 in 2010 to a minimum of 0.157 in 2014, subsequently rising to 0.185 by 2024. This pattern suggests that North Dakota's relative attractiveness to international migrants has recovered following the mid-decade trough, potentially reflecting post-oil-boom economic restructuring and the resumption of refugee resettlement programs.

3.2 Geographic Concentration

Analysis of geographic concentration reveals a divergence between country-level and regional-level patterns. The Herfindahl-Hirschman Index (HHI) for country-of-origin composition is 1,162 based on FY 2023 LPR admissions, classifying North Dakota's immigrant origins as "unconcentrated" under standard thresholds (HHI $< 1,500$). This result indicates substantial diversity at the country level, with no single origin dominating the flow.

However, regional aggregation yields markedly different conclusions. The regional-level HHI reaches 3,712, placing North Dakota firmly in the "highly concentrated" category (HHI $> 2,500$). This concentration derives from the dominance of two source regions: Asia accounts for 52.5% of LPR admissions and Africa accounts for 27.9%, together comprising over 80% of legal permanent resident flows. North America (12.3%), Europe (3.4%), South America (3.4%), and Oceania (0.6%) contribute the remainder.

Location Quotient (LQ) analysis identifies countries substantially overrepresented in North Dakota relative to national immigrant composition. Table 5 reports LQ values for the ten most overrepresented origins. Egypt exhibits the highest LQ (15.13), indicating that North Dakota's share of Egyptian-born residents exceeds the national share by a factor of fifteen. Other highly

overrepresented origins include India (LQ = 9.86), Sudan (LQ = 8.21), and several African regions. This pattern reflects the legacy of refugee resettlement programs that have directed migrants from specific origin countries to North Dakota.

Table 5: Location Quotients: Top Overrepresented Origin Countries, 2023

Country	ND Foreign-Born	ND Share (%)	US Share (%)	LQ
Egypt	373	1.07	0.07	15.13
India	329	0.95	0.10	9.86
Other Western Africa	400	1.15	0.12	9.23
Sudan	164	0.47	0.06	8.21
Other Middle Africa	903	2.59	0.36	7.26
Pakistan	860	2.47	0.37	6.63
Croatia	526	1.51	0.23	6.56
Other Eastern Africa	3,582	10.29	1.66	6.18
Kenya	11,920	34.24	5.77	5.94
South Africa	387	1.11	0.19	5.93

Notes: LQ = Location Quotient. Values exceeding 1.0 indicate overrepresentation relative to national composition. Data from American Community Survey 2023.

The concentration patterns exhibit meaningful temporal evolution. Country-level HHI has increased from 482 in 2012 to 942 in 2023, though the trend is not statistically significant ($p = 0.242$). Regional-level HHI shows more pronounced increases, rising from approximately 2,600 in the early 2010s to over 5,200 in 2022–2023, reflecting growing reliance on African and Asian source regions.

3.3 Time Series Properties

Unit root diagnostics inform integration choices for subsequent time series modeling. Table 6 reports Augmented Dickey-Fuller (ADF) test results for the three primary series.

Table 6: Unit Root Test Results

Variable	ADF (Level)	p-value	ADF (Diff)	p-value
ND International Migration	−1.453	0.556	−3.843	0.002
ND Share of US	−2.004	0.285	−6.214	< 0.001
US International Migration	−3.908	0.002	−2.347	0.157
Integration Order				
ND International Migration	I(1) — First-difference stationary			
ND Share of US	I(1) — First-difference stationary			
US International Migration	I(0) — Level stationary			

Notes: ADF = Augmented Dickey-Fuller test statistic. Lag selection by AIC. $n = 15$ annual observations. Critical values at 5%: −3.29 (level), −3.13 (differenced).

North Dakota international migration does not reject the unit root null in levels ($p = 0.556$)

but rejects after first differencing ($p = 0.002$), consistent with I(1) behavior. North Dakota's share of national migration shows the same ADF pattern (level $p = 0.285$; differenced $p < 0.001$). In contrast, national international migration rejects the unit root null in levels ($p = 0.002$), consistent with I(0) behavior.

KPSS tests (null: level-stationarity; regression = constant) do not reject level-stationarity for the ND migration series ($p \geq 0.10$), while KPSS rejects level-stationarity for the first-differenced ND share series ($p \approx 0.042$). Together with ADF results, the small-sample diagnostics are suggestive rather than definitive: integration order is treated as a modeling choice, with break-robust checks (Appendix B.1) used to evaluate sensitivity.[3]

ARIMA model selection via AIC identifies ARIMA(0,1,0) as a candidate baseline specification in this very small sample. The random-walk baseline implies that the best one-step forecast equals the most recent observation, with innovations representing unpredictable shocks. Diagnostic tests support adequacy within the limited sample: the Ljung-Box portmanteau test finds no significant residual autocorrelation through lag 5 ($Q = 4.12$, $p = 0.404$), and residuals pass the Shapiro-Wilk normality test ($W = 0.968$, $p = 0.820$).

The random-walk baseline implies substantial forecast uncertainty. Five-year-ahead prediction intervals widen dramatically: the 2029 forecast of 5,126 migrants carries a 95% prediction interval of [212, 10,040], spanning nearly two orders of magnitude. This uncertainty reflects both the inherent unpredictability of migration processes and the limited sample size constraining parameter precision.

Rolling-origin backtesting (expanding window, 2010–2016 start; 2017–2024 forecasts) benchmarks point accuracy and interval calibration. Table 7 reports Mean Absolute Scaled Error (MASE) alongside conventional metrics; MASE is robust to the near-zero 2020 observation that inflates MAPE. The driver regression using contemporaneous U.S. migration—an "oracle" benchmark that assumes knowledge of future national flows—achieves the lowest MAE/RMSE and MASE = 0.27. However, this model is infeasible for real-time forecasting. A feasible variant using lagged (prior-year) U.S. migration achieves MASE = 1.06, essentially matching the naive baseline. Among all feasible methods, the naive random-walk and ARIMA(0,1,0) perform identically (MASE = 1.00), with no feasible model significantly outperforming the random walk. Interval widths remain large across models; these results are therefore descriptive rather than definitive.

Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots for the differenced series (Figure 4) confirm the absence of significant serial correlation at standard lags, consistent with the random-walk baseline.

3.4 Structural Break Analysis

Structural break tests evaluate whether regression parameters remain stable across the observation period or exhibit discrete shifts at specific dates. Table 8 reports Chow test results for three candidate break points corresponding to major policy events: the 2017 Travel Ban, the 2020 COVID-19

[3]Statsmodels reports KPSS p-values as 0.10 when $p \geq 0.10$; values shown as 0.10 should be read as upper bounds.

Table 7: Rolling-Origin Backtesting Results (h = 1)

Model	MAE	RMSE	MASE	MAPE	Cov. 80%	Cov. 95%
Feasible methods (use only past data):						
Naive RW	1,153	1,367	1.00	313%	0.63	0.75
Expanding mean	1,674	1,980	1.45	635%	0.38	0.50
Driver OLS (Lagged)	1,224	1,408	1.06	534%	0.38	0.63
ARIMA(0,1,0)	1,153	1,367	1.00	313%	0.63	0.75
Infeasible benchmark (requires future data):						
Driver OLS (Oracle)†	317	530	0.27	65%	0.88	0.88

Notes: Rolling-origin evaluation with expanding window; $n = 8$ one-step forecasts. MASE = Mean Absolute Scaled Error (scaled by naive RW MAE; values <1 outperform naive baseline). Coverage refers to empirical coverage of symmetric prediction intervals. MAPE is inflated by the near-zero 2020 observation; MASE provides a more robust accuracy comparison.
†*Oracle benchmark*: Driver OLS with contemporaneous U.S. migration values, unavailable at forecast time. The lagged driver—a feasible variant using prior-year U.S. migration—achieves MASE = 1.06, essentially matching the naive baseline. No feasible model significantly outperforms random walk.

pandemic, and the 2021 post-pandemic recovery.

Table 8: Structural Break Test Results (Chow Tests)

Break Year	Policy Event	F-statistic	p-value	Regime Shift
2017	Travel Ban	1.29	0.314	Not significant
2020	COVID-19	16.01	< 0.001	+91.1%
2021	Post-COVID Recovery	10.28	0.003	+161.6%

Notes: Chow test with df = (2, 11). Critical values: 3.98 (5%), 7.21 (1%). Regime shift indicates percentage change in mean migration between pre- and post-break periods.

The 2020 COVID-19 pandemic represents a highly significant structural break ($F = 16.01$, $p < 0.001$). Mean international migration to North Dakota increased from 1,378 in the pre-2020 period to 2,633 in the post-2020 period—a 91.1% increase. However, this result requires careful interpretation: the 2020 observation itself (30 migrants) represents partial-year data during severe travel restrictions, while subsequent years (2021–2024) reflect both recovery and potentially accelerated growth. The break thus captures a shift to a higher-volatility regime rather than a simple level increase.

The 2021 break point similarly proves significant ($F = 10.28$, $p = 0.003$), with mean migration increasing from 1,255 (2010–2020) to 3,284 (2021–2024)—a 161.6% change. This breakpoint more cleanly separates the COVID-affected period from the subsequent recovery.

Notably, the 2017 Travel Ban shows no statistically significant break at the aggregate North Dakota level ($F = 1.29$, $p = 0.314$). This null result contrasts with the significant Travel Ban effects identified in the nationality-level difference-in-differences analysis (Section 3.7), suggesting that aggregate flows may mask compositional shifts in origin countries.

The CUSUM test for parameter stability corroborates these findings. The cumulative sum of recursive residuals remains within the 5% significance bounds throughout the observation period (maximum CUSUM = 2.37, bound = 10.25), indicating no evidence of continuous parameter drift. The stability ratio of 0.23 suggests that cumulative deviations from the fitted model remain modest relative to permissible fluctuation.

The Bai-Perron procedure for endogenous break detection identifies zero breaks under the BIC criterion, reflecting the penalty imposed on additional parameters in the short series. This result highlights the tension between data-driven and theory-driven break identification in small samples.

3.5 Panel Data Results

Panel analysis exploits cross-state and temporal variation to examine determinants of international migration allocation. The estimation sample comprises 765 state-year observations (51 states $\times$ 15 years), forming a balanced panel.

The Hausman specification test compares fixed effects and random effects estimators. Under the null hypothesis that random effects estimates are consistent and efficient, the test statistic follows a chi-squared distribution. The computed test statistic is effectively zero ($H = 7.18 \times 10^{-31}$, $p = 1.000$), indicating no systematic difference between estimators and thus supporting random effects as the preferred specification. The Breusch-Pagan LM test strongly rejects pooled OLS in favor of the random effects model ($\chi^2 = 1,502$, $p < 0.001$), confirming the presence of state-level heterogeneity requiring panel methods.

State fixed effects from the two-way specification reveal substantial cross-sectional variation in international migration levels. Florida exhibits the largest positive effect (+125,526 relative to the grand mean), followed by California (+106,259) and Texas (+89,232). North Dakota's fixed effect of −18,022 places it near the bottom of the distribution, exceeded in negativity only by Wyoming (−19,332). This ranking accords with population-weighted expectations and confirms North Dakota's status as a minor destination in absolute terms.

Year fixed effects capture common temporal shocks affecting all states. The 2020 effect is dramatically negative, reflecting the near-cessation of international migration during pandemic travel restrictions. Mean international migration in 2020 was 390 compared to 21,206 in other years—a 98.2% reduction. The between-within variance ratio of 1.30 indicates that cross-state variation slightly exceeds within-state temporal variation, though both dimensions contribute meaningfully to total variability.

3.6 Cross-Sectional Allocation Model Estimates

Cross-sectional allocation models estimated via Poisson pseudo-maximum likelihood (PPML) quantify the relationship between LPR admissions and diaspora stocks. As noted in Section 2.6, these models differ from traditional gravity specifications in that they omit origin–destination distance, which varies minimally across U.S. states for a given origin country. Table 9 reports coefficient estimates across three specifications of increasing complexity.

Table 9: Cross-Sectional Allocation Model Estimates (PPML)

Variable	**Simple**	**Full Model**	**State FE**
Log(Diaspora Stock)	0.448***	0.140	0.163
	(0.104)	(0.248)	(0.148)
Log(Origin Stock in U.S.)	—	0.046	—
		(0.192)	
Log(State Foreign-Born Total)	—	0.792***	—
		(0.212)	
Constant	1.267	−7.711*	—
	(0.980)	(3.123)	
n	4,845	4,845	4,845
Pseudo R^2	0.236	0.399	0.413

Notes: Two-way clustered standard errors (state and origin) in parentheses. PPML estimation. * $p < 0.05$; *** $p < 0.001$. Dependent variable: LPR admissions by state-country pair, FY 2023. Distance is omitted because within-U.S. variation in origin–state distance is minimal (see Section 2.6). State FE specification includes 50 state dummy variables (coefficients not reported).

The simple specification yields a diaspora association of 0.448 (SE = 0.104, $p < 0.001$), indicating that larger state–origin stocks are correlated with higher FY2023 admissions. However, this bivariate model omits origin and destination scale.

The full model adds origin and destination foreign-born totals. With two-way clustered SEs, the diaspora association attenuates to 0.140 (SE = 0.248, $p = 0.57$), while state foreign-born totals remain the dominant predictor (elasticity = 0.792, SE = 0.212, $p < 0.001$). The origin stock coefficient is small and imprecise (0.046, SE = 0.192).

The state fixed effects specification yields a diaspora association of 0.163 (SE = 0.148, $p = 0.27$). This estimate remains below the simple model and underscores the sensitivity of cross-sectional coefficients to controls and inference.

The attenuation from 0.45 to approximately 0.14 has practical implications for forecasting. Much of the simple-model association reflects selection of large receiving states with both large diaspora stocks and continued large flows. The pseudo-R^2 values indicate that the full model explains 40% of deviance, roughly double the explanatory power of the bivariate specification.

Because FY2023 is a single cross-section, these results are interpreted as predictive diaspora–flow associations rather than causal diaspora effects. We report two-way clustered SEs to account for correlation within origins and destinations, and note that ACS diaspora stocks are estimates with margins of error; coefficient uncertainty does not fully propagate that measurement error.

3.7 Causal Inference Findings

Causal inference results focus on difference-in-differences for the Travel Ban and an interrupted time series for COVID-19. A synthetic comparator is reported as a descriptive benchmark, and a shift-share index provides first-stage relevance for refugee-driven shocks.

Table 10: Policy Effects and Predictive Relevance

Method	Policy	Estimate	SE	95% CI
DiD (log scale)	Travel Ban	-1.384*	0.646	$[-2.65, -0.12]$
Percentage effect		-74.9%	—	—
ITS Level Shift	COVID-19	$-19,503$***	4,556	$[-28,432, -10,573]$
ITS Trend Change	COVID-19	$+14,113$***	3,288	$[+7,667, +20,558]$
Shift-share (first stage)	National shocks	5.51***	0.87	$[+3.80, +7.22]$
First-stage F		39.95		

Notes: DiD = Difference-in-differences with nationality and year fixed effects, SEs clustered by nationality. ITS = Interrupted time series with state fixed effects, SEs clustered by state. Shift-share row reports first-stage relevance only (not a causal effect). * $p < 0.05$; *** $p < 0.001$. Travel Ban affected countries: Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen.

3.7.1 Travel Ban Difference-in-Differences

The Travel Ban analysis employs refugee arrival data spanning 2002–2019, with 2018 designated as the first full post-treatment year. The treatment group comprises seven affected nationalities (Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen), while 119 unaffected nationalities serve as controls. The sample includes 1,137 nationality-year observations.

The estimated average treatment effect on the treated (ATT) is -1.384 on the log scale ($t = -2.14$, $p = 0.032$), corresponding to a policy-associated divergence of 74.9% in refugee arrivals from affected countries. The 95% confidence interval $[-2.65, -0.12]$ excludes zero, indicating a statistically significant and economically substantial short-run divergence.

Parallel trends diagnostics are mixed, requiring caution in causal interpretation. The linear pre-trend test yields a treated-country trend coefficient of 0.087 ($t = 0.97$, $p = 0.334$), but the event study's joint pre-treatment test rejects parallel trends over the full pre-period ($F = 4.31$, $p < 0.001$). Figure 7 shows divergence beginning approximately 10 years before treatment, indicating that the treatment and control groups were on different trajectories prior to the policy. The immediate post-treatment coefficients shift sharply negative in 2018–2019, but because the parallel trends assumption is violated, the estimated effect conflates the policy impact with pre-existing divergence. We therefore interpret the 75% estimate as an upper bound on the causal policy effect rather than a point estimate of the true effect.

Small-sample inference robustness. With only seven treated clusters (nationalities), conventional cluster-robust inference may be unreliable. Table 11 reports alternative p-values from wild cluster bootstrap and randomization inference. For the full pre-period specification, the wild cluster bootstrap yields $p = 0.077$ and randomization inference yields $p = 0.058$—both above the conventional $\alpha = 0.05$ threshold, though significant at $\alpha = 0.10$. This attenuation relative to the conventional $p = 0.032$ reflects the small number of treated clusters and suggests caution in claiming strong statistical significance.

Restricted pre-period robustness. Because the event study reveals greater pre-trend diver-

Table 11: Small-Sample Inference Robustness for Travel Ban DiD

Specification	**ATT**	**Conventional** p	**WCB** p	**RI** p
Full pre-period (2002–2017)	−1.384	0.032*	0.077	0.058
Restricted (2013–2017)	−1.983	<0.001***	0.003**	0.004**

Notes: WCB = Wild cluster bootstrap with Rademacher weights (999 iterations). RI = Randomization inference with treatment permutation (999 permutations). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

gence in early years (2002–2012) than in years immediately preceding treatment, we re-estimate the DiD restricting to years where parallel trends appear more plausible. Restricting the pre-period to 2013–2017 substantially improves the pre-trend diagnostic: the joint F-test yields $F = 0.76$, $p = 0.553$, failing to reject parallel trends. The ATT estimate increases in magnitude to −1.983 (SE = 0.506), and inference strengthens: the conventional p-value falls to $p < 0.001$, with bootstrap ($p = 0.003$) and randomization ($p = 0.004$) inference confirming significance at the 1% level.

This pattern—larger ATT magnitude and stronger significance with the restricted pre-period—suggests that the full-sample estimate is attenuated by pre-existing negative divergence between treatment and control groups. When analysis focuses on years where parallel trends are more defensible, the estimated policy effect is larger (−86% vs. −75%) and more precisely estimated. The restricted-period results strengthen confidence that the Travel Ban produced a genuine reduction in refugee arrivals from affected countries, though the magnitude remains subject to specification uncertainty.

3.7.2 COVID-19 Interrupted Time Series

The COVID-19 analysis employs state-level panel data (2010–2024) with 2020 as the intervention year. The interrupted time series model identifies both level and trend effects.

Scope of inference: The ITS specification includes state fixed effects and estimates the *average* COVID disruption across all U.S. states, not a North Dakota-specific effect. Because COVID-19 and associated travel restrictions affected all states simultaneously, the model cannot separately identify whether North Dakota experienced a differential disruption relative to other states. The coefficients reported below describe the mean state-level response; ND-specific deviations from this average are absorbed by the state fixed effect and are not separately estimated. Extending the model to include ND-specific interaction terms (e.g., ND × Post_{2020}) would require additional identifying assumptions and is left for future work.

The level shift coefficient indicates that international migration dropped by 19,503 persons immediately following COVID onset ($t = -4.28$, $p < 0.001$). The trend change coefficient of +14,113 ($t = 4.29$, $p < 0.001$) reflects the steep post-pandemic rebound slope during 2021–2024. Relative to pre-COVID mean migration of 15,669, the level shift represents a 124.5% decline—interpretable as a transition from positive to negative net migration during the acute pandemic period.

Interpretive caveat: The large positive trend change coefficient is a mathematical artifact of fitting a linear trend to the short, steep recovery period (2021–2024). With pre-COVID mean migration of approximately 1,800 persons annually, a coefficient of +14,113 would imply explosive growth if extrapolated—clearly implausible. This coefficient should be interpreted as describing the short-run "rebound slope" rather than a sustainable long-term trajectory. Extrapolation beyond the observed post-COVID window is not warranted.

3.7.3 Synthetic Comparator (Descriptive Benchmark)

Because the Travel Ban is a national shock, we do not interpret a state-level synthetic comparator as causal. Instead, we construct a synthetic comparator that matches North Dakota's pre-2017 international migration rate using weighted donor states. The weights emphasize Wyoming (41.9%), Vermont (24.5%), Rhode Island (20.1%), and Washington (8.3%), with minor contributions from Oregon (2.7%) and Florida (2.5%).

Pre-treatment fit is strong: the root mean squared prediction error (RMSPE) of 0.020 indicates close tracking during 2010–2016. The post-2017 gap averages +0.57 migrants per 1,000 population (SD = 0.78), and the post/pre RMSPE ratio is large (48.6), indicating substantial divergence. These deviations are presented descriptively and are not interpreted as policy effects.

3.7.4 Shift-Share (Bartik) Index

The shift-share index combines baseline state shares with leave-one-out national origin shocks to summarize predicted refugee inflows. The first-stage F-statistic of 39.95 exceeds the conventional threshold of 10, supporting instrument relevance.

The estimated first-stage coefficient is 5.51 ($p < 0.001$), indicating that a one-unit increase in the shift-share index predicts a 5.51-unit increase in state-level international migration. Because no second-stage causal model is estimated, this result is interpreted as predictive relevance rather than a causal effect.

3.8 Duration Analysis

Duration analysis examines the lifecycle of immigration "waves"—sustained periods of elevated arrivals from specific origin countries. Wave identification criteria require arrivals to exceed 150% of baseline for at least two consecutive years. This procedure identifies 940 waves across 56 nationalities and 48 states in the 2002–2020 refugee data.

Kaplan-Meier estimation yields a median wave duration of 3.0 years (mean = 3.54 years). Survival probabilities decline steadily over the wave lifecycle: 51.6% of waves survive past year 2; 31.8% survive past year 3; 21.6% survive past year 4; and only 6.2% survive past year 10. The censoring rate of 10% reflects waves still active at the end of the observation period (FY 2020).

Stratified analysis reveals significant heterogeneity by wave intensity. Table 12 reports median survival by intensity quartile.

Table 12: Wave Duration by Intensity Quartile

Intensity Quartile	n	**Median Survival**	**Mean Duration**
Q1 (Low)	247	2.0 years	2.43 years
Q2	223	2.0 years	2.80 years
Q3	236	3.0 years	3.61 years
Q4 (High)	234	4.0 years	5.35 years
Log-rank test: $\chi^2 = 278.7$, $p < 10^{-60}$			

Notes: Intensity = peak arrivals relative to baseline. Duration measured in years. Log-rank test evaluates equality of survival curves across quartiles.

The log-rank test decisively rejects equality of survival curves across intensity quartiles ($\chi^2 = 278.7$, $p < 10^{-60}$), confirming that higher-intensity waves persist substantially longer. The median survival differential between Q1 and Q4—two years versus four years—indicates that wave duration doubles at the upper extreme of initial intensity.

Regional origin also influences wave persistence. African (mean duration = 3.99 years), Middle Eastern (4.14 years), and Asian (4.28 years) waves outlast those from Europe (2.40 years) and the Americas (2.25 years). The log-rank test confirms significant regional heterogeneity ($\chi^2 = 89.7$, $p < 10^{-17}$).

Cox proportional hazards regression quantifies the multivariate relationship between wave characteristics and termination risk. Key findings include:

- **Log intensity**: HR = 0.412 ($p < 0.001$). Higher-intensity waves exhibit 59% lower hazard of termination, confirming the protective effect of initial magnitude.
- **Early wave**: HR = 1.361 ($p < 0.001$). Waves beginning in the early observation period terminate faster, potentially reflecting data truncation or secular trends.
- **Peak arrivals**: HR = 0.656 ($p < 0.001$). Higher peak arrivals during the wave reduce termination hazard.
- **Americas origin**: HR = 1.711 ($p = 0.005$). Waves from the Americas terminate 71% faster than the reference category (Africa).
- **Europe origin**: HR = 1.568 ($p = 0.004$). European-origin waves similarly exhibit elevated termination hazard.

The Cox model achieves a concordance index of 0.769, indicating good discriminative ability—the model correctly ranks wave pairs by survival time 77% of the time. The proportional hazards assumption is satisfied based on the global Schoenfeld residual test.

To connect duration estimates to forecasting, we translate the Cox model into year-ahead survival probabilities for any active wave. For a wave observed at age a, the conditional survival ratio $S(a + k \mid x)/S(a \mid x)$ yields the probability the wave persists k additional years, which we combine

with the empirically estimated lifecycle shape (initiation → peak → decline) to simulate wave contributions within the Monte Carlo scenario engine. Figure 8 summarizes the survival curves that anchor these persistence draws.

3.9 Forecast Scenarios

Scenario analysis translates analytical findings into forward-looking projections through 2045. Four policy-indexed scenarios span the range of plausible trajectories, while Monte Carlo simulation quantifies parametric uncertainty. Although the Driver OLS model achieves superior backtesting accuracy (Table 7), it requires contemporaneous knowledge of national migration—information unavailable at forecast time. Long-term scenarios therefore employ the conservative ARIMA baseline, which relies only on historical data and represents the best-performing feasible univariate method. Table 13 summarizes 2045 projections under each scenario.

Table 13: Net International Migration (PEP) Scenario Projections for North Dakota, 2045

Scenario	**Assumptions**	**2045 Projection**
CBO Full	2025–2029: 1.1×ARIMA; 2030–2045: 8% growth	19,318
Moderate	Dampened historical trend (50% weight)	7,048
Pre-2020 Trend	Anchor 2019; continue 2010–2019 slope (+72/year)	2,517
Zero	No international migration	0
Monte Carlo Median	Stochastic trend with $CV_{2045} = 0.38$	9,056
50% PI		[6,450, 11,292]
95% PI		[3,570, 14,491]

Notes: CBO = Congressional Budget Office elevated immigration scenario. PI = prediction interval from 1,000 Monte Carlo draws. CV_{2045} denotes the coefficient of variation of simulated 2045 outcomes (historical CV is 82.5%). Baseline 2024 value = 5,126.

The four deterministic scenarios yield dramatically different endpoints. The CBO Full scenario applies 10% above-ARIMA levels for 2025–2029 and then compounds at 8% annually, projecting net international migration of 19,318 persons by 2045. The Moderate scenario, applying 50% dampening to historical trends, yields 7,048. The Pre-2020 Trend scenario anchors at the 2019 level and extrapolates the 2010–2019 slope, projecting only 2,517—reflecting the modest pre-pandemic growth rate of approximately 72 persons per year.

Monte Carlo simulation propagates parameter uncertainty through the projection model. Based on 1,000 draws from estimated trend and volatility distributions ($CV_{2045} = 0.38$), the median 2045 projection is 9,056 persons, with a 50% prediction interval of [6,450, 11,292] and a 95% prediction interval of [3,570, 14,491]. The latter interval spans a factor of 4.1, reflecting substantial irreducible uncertainty over a 20-year horizon given the observed volatility and structural instability. The simulation CV refers to dispersion in projected 2045 outcomes and is not directly comparable to the historical CV of annual observations.

Model averaging weights alternative time series specifications by AIC. The VAR model receives

effectively all AIC weight (1.000) relative to the univariate ARIMA (weight ≈ 0), reflecting the superior fit of the multivariate specification. For cross-sectional models, Random Forest receives the highest R^2-based weight (0.462), followed by Elastic Net (0.293) and OLS (0.245).

Role of Machine Learning Methods. The ML methods described in Section 2.7 serve an auxiliary role in this analysis rather than contributing directly to the scenario trajectories. Elastic Net regularization ($R^2_{CV} = 0.60$) identifies three non-zero predictors of state-level international migration share: log population (coefficient = 0.015), natural change (0.011), and birth rate (-0.003). Random Forest ($R^2_{CV} = 0.94$) confirms that log population dominates feature importance (93% of relative importance), with death rate and domestic migration rate contributing marginally. K-means clustering ($K = 2$; silhouette = 0.64) places North Dakota in the larger cluster containing 52 states, with only one state forming a distinct cluster.

These ML results provide three contributions to the broader analysis: (1) they validate that population scale is the primary determinant of state-level migration allocation, consistent with the cross-sectional models in Section 3.6; (2) they identify peer states for North Dakota (the 51 states in its cluster) useful for comparative benchmarking; and (3) they confirm that no parsimonious feature set substantially improves upon simple population-based allocation models. Because ML methods are trained on cross-sectional variation and do not generate time-series forecasts, they do not directly parameterize the scenario engine. Full ML diagnostics are available in the supplementary materials.

The fan chart visualization (Figure 9) depicts the scenario trajectories with Monte Carlo uncertainty bands, illustrating both the range of policy-contingent outcomes and the widening uncertainty envelope over the projection horizon.

4 Discussion

This study set out to characterize international migration to North Dakota—a small state operating at the demographic margins of American immigration—and to develop a rigorous framework for forecasting these flows under uncertainty. The multi-method analysis yielded convergent findings across nine analytical modules, while also revealing the inherent limits of prediction in this domain. This section interprets the key findings in relation to the research questions posed in the introduction, situates them within the broader migration literature, considers policy implications, and acknowledges limitations.

4.1 Interpretation of Key Findings

4.1.1 Patterns and Sources of International Migration

The geographic concentration analysis reveals a distinctive migration profile that diverges markedly from national patterns. While North Dakota's country-level immigrant origins appear diverse (HHI

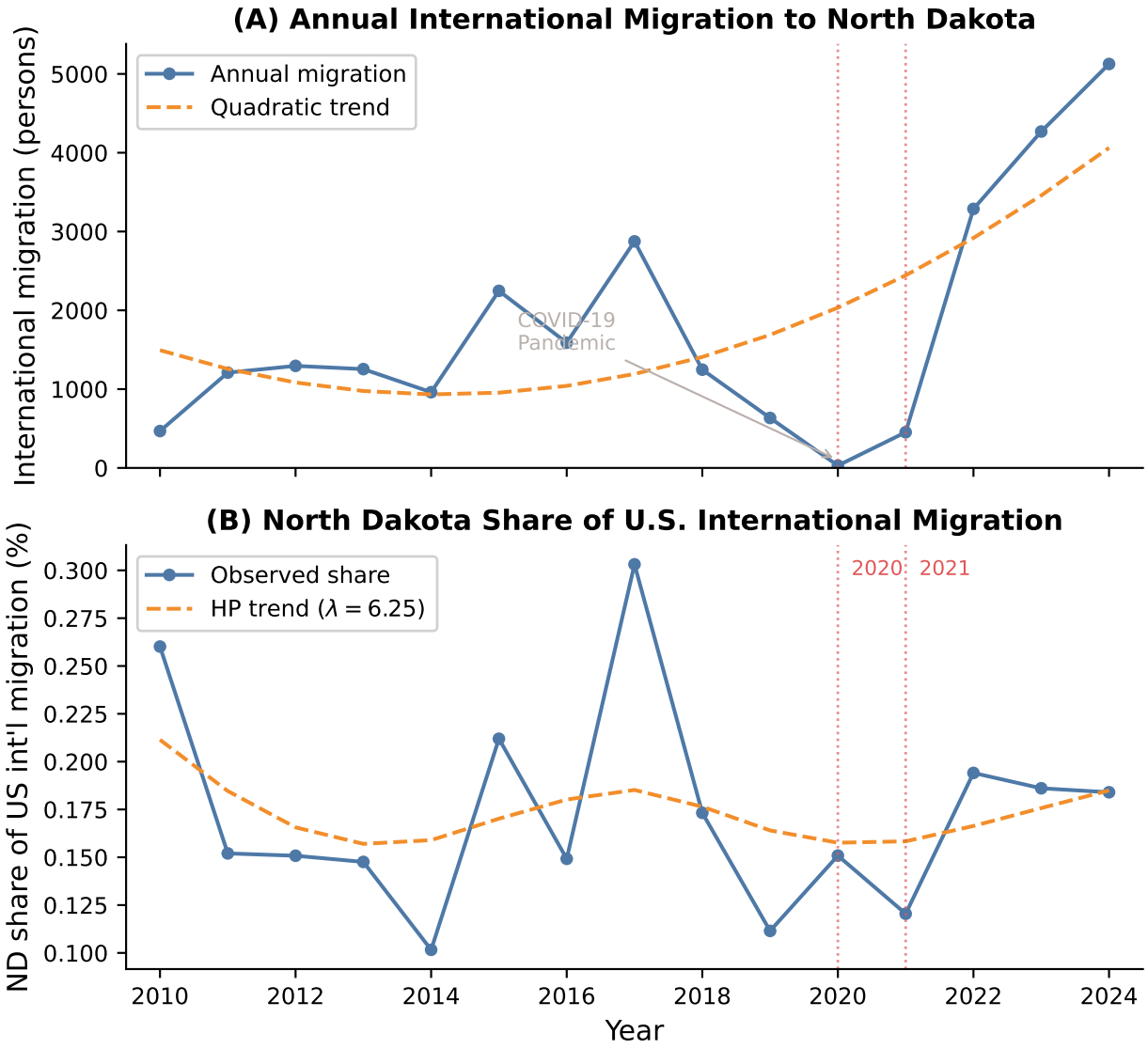

Figure 2: International migration to North Dakota, 2010–2024. Panel (A) shows annual migration flows with a quadratic trend line. Panel (B) displays North Dakota's share of total U.S. international migration, with the solid line showing observed values and the dashed line indicating the HP filter trend component ($\lambda = 6.25$). Vertical dotted lines mark identified structural breaks at 2020 (COVID-19 pandemic) and 2021 (recovery period), both significant at $p < 0.01$ using Chow tests.

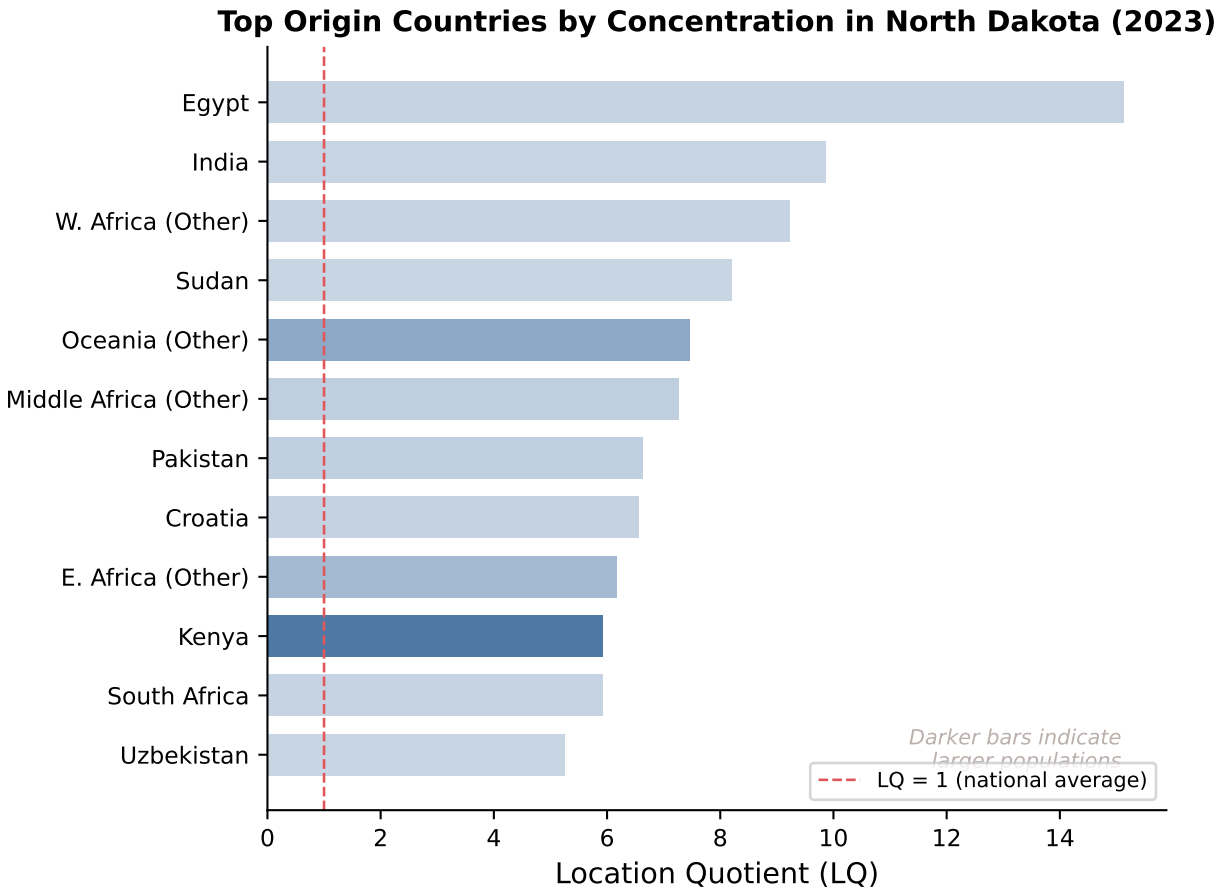

Figure 3: Geographic concentration of foreign-born population in North Dakota by country of origin (2023). Bars show location quotients (LQ), where LQ > 1 indicates overrepresentation relative to the national average. The dashed vertical line marks LQ $= 1$. Countries are sorted by LQ value; bar intensity reflects population size (darker indicates larger foreign-born population in North Dakota). Egypt, India, and several African origin countries show the highest concentration relative to their national presence.

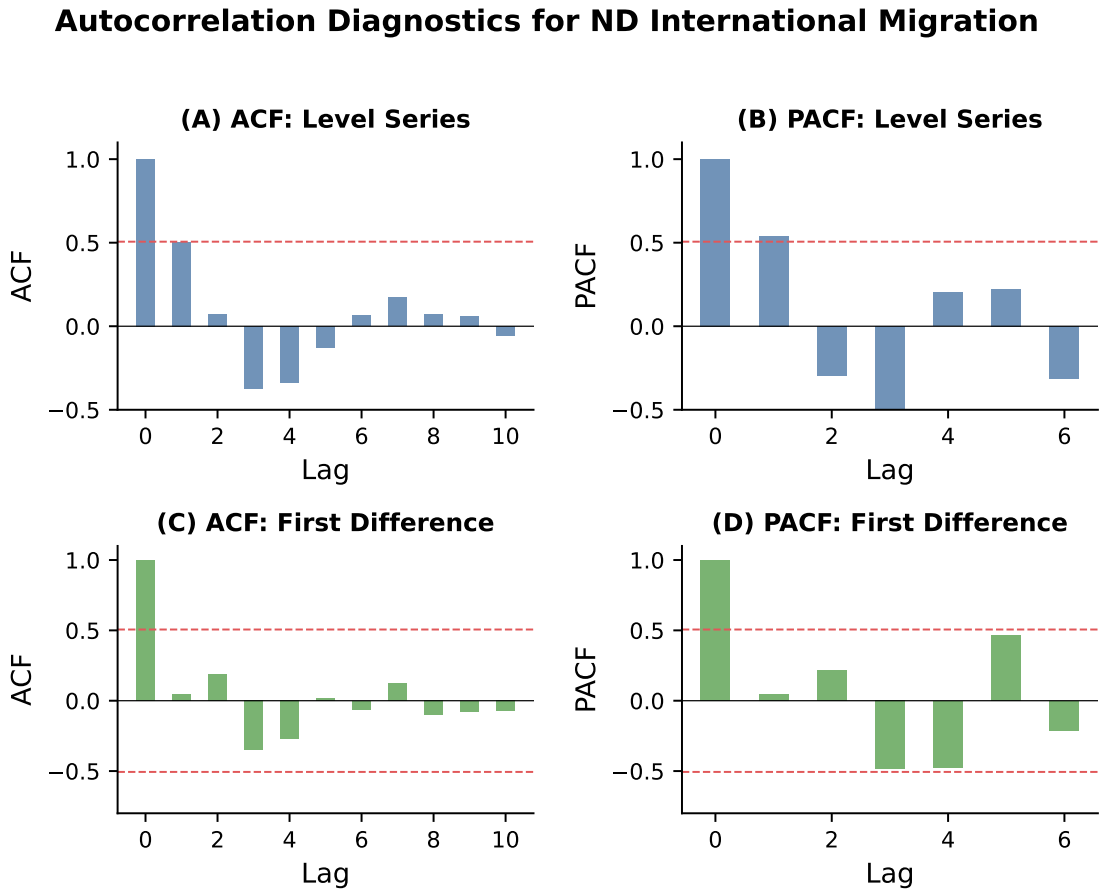

Figure 4: Autocorrelation function (ACF) and partial autocorrelation function (PACF) for North Dakota international migration series. Panels (A) and (B) show the level series, which exhibits slow decay in the ACF consistent with nonstationarity. Panels (C) and (D) show the first-differenced series, where autocorrelations fall within the 95% confidence bounds (dashed lines), supporting the characterization of the series as I(1). Sample size $n = 15$.

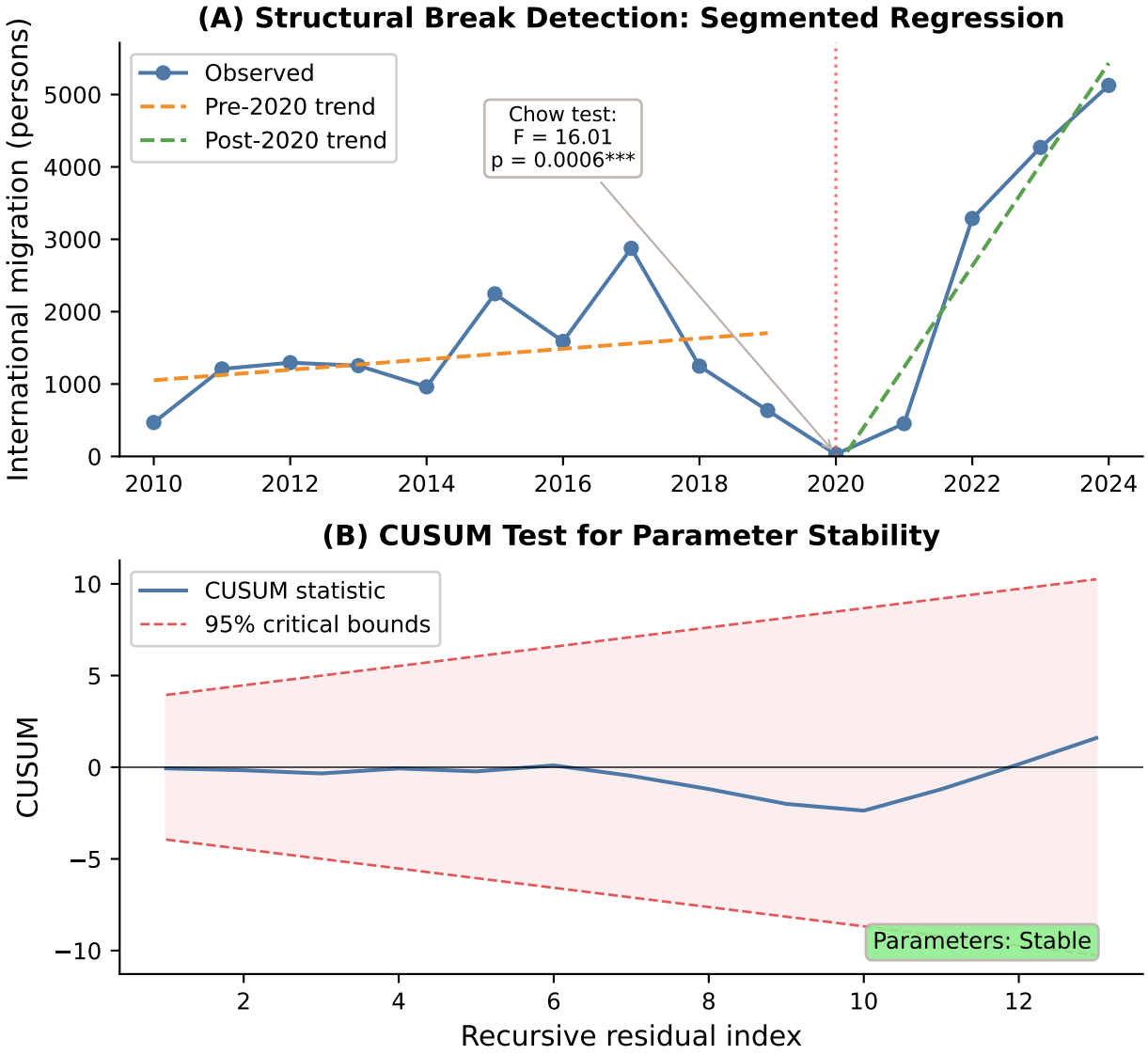

Figure 5: Structural break analysis for North Dakota international migration. Panel (A) shows the time series with segmented regression fits before and after 2020, with Chow test statistics indicating a significant break ($F = 16.01$, $p < 0.001$). Panel (B) displays the CUSUM test for parameter stability, where the test statistic (solid line) remains within the 95% critical bounds (dashed lines), suggesting overall parameter stability despite the level shift.

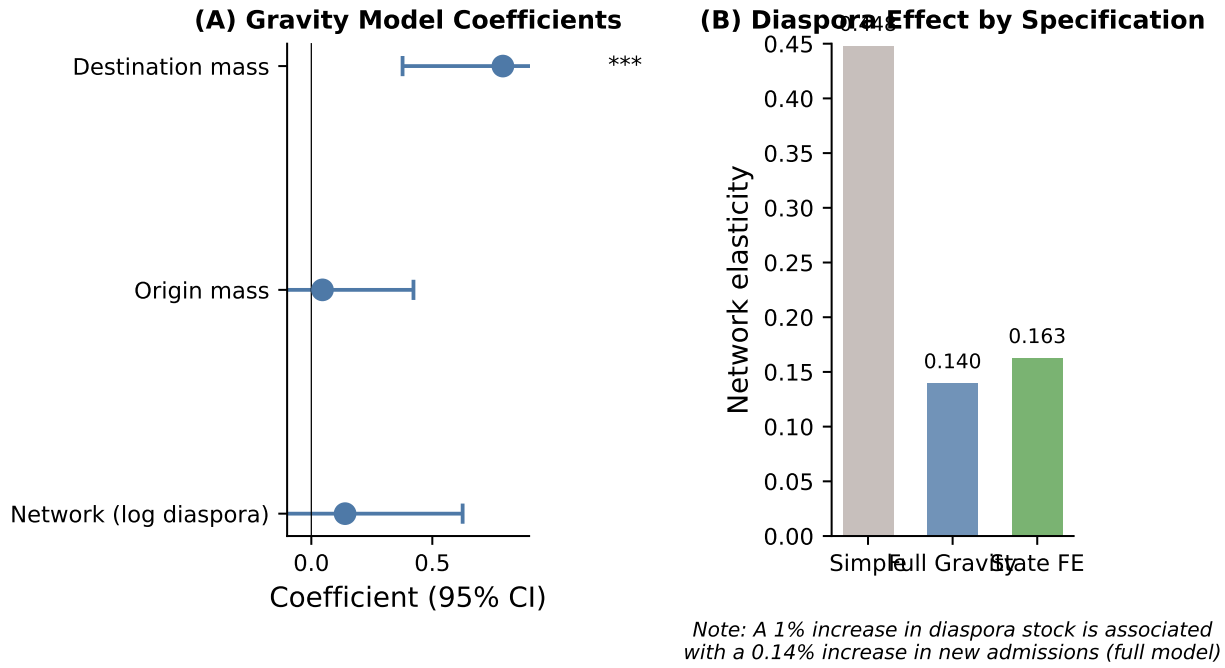

Figure 6: Gravity model estimation results. Panel (A) shows coefficient estimates with 95% confidence intervals for the full gravity specification: diaspora association (log diaspora stock), origin mass (log origin stock in U.S.), and destination mass (log state foreign-born total). Panel (B) compares diaspora association estimates across model specifications, showing that controlling for mass variables reduces the diaspora coefficient from 0.45 to approximately 0.14.

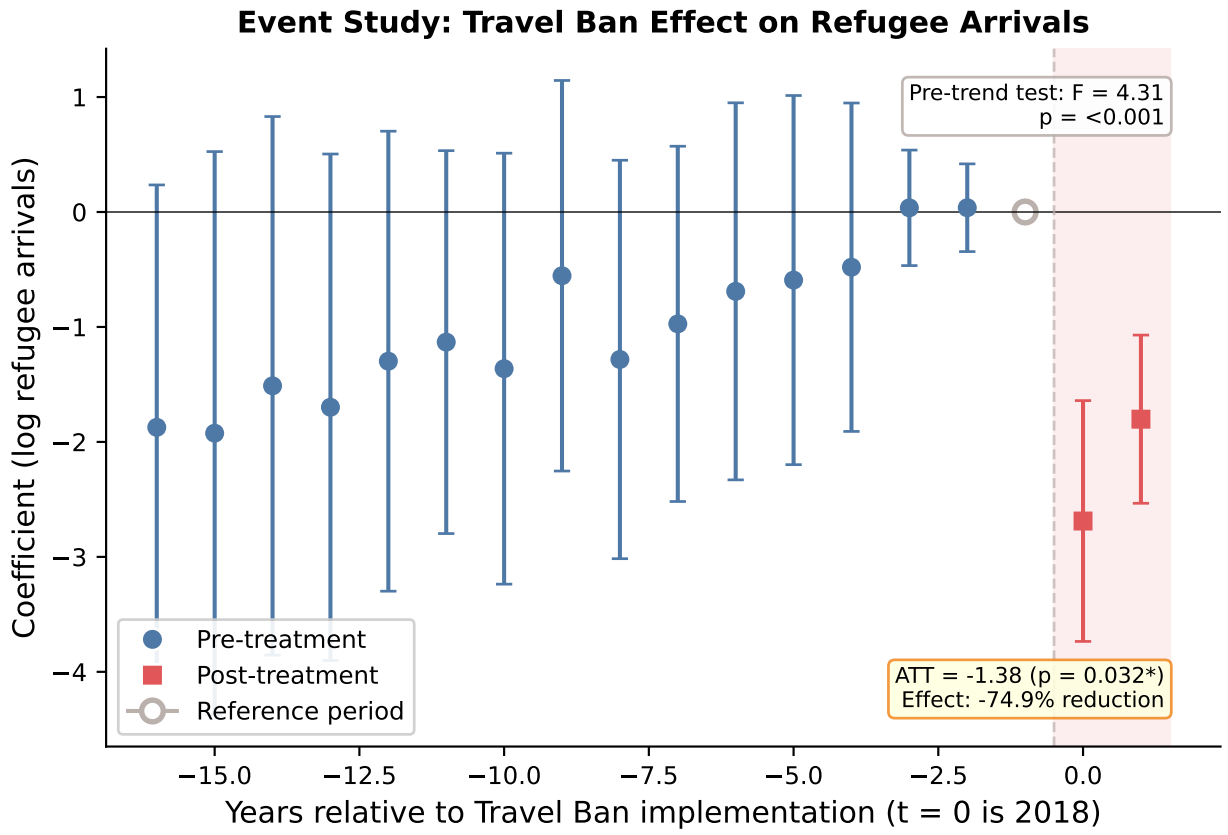

Figure 7: Event study estimates for the Travel Ban effect on refugee arrivals. Coefficients represent the difference in log arrivals between treated countries (Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen) and control countries relative to the reference period ($t = -1$, 2017). Blue circles show pre-treatment estimates; red squares show post-treatment effects. The joint pre-trend test rejects parallel trends over the full pre-period ($F = 4.31$, $p < 0.001$). The average treatment effect on the treated (ATT) is -1.38 ($p = 0.032$), corresponding to a 74.9% reduction in arrivals.

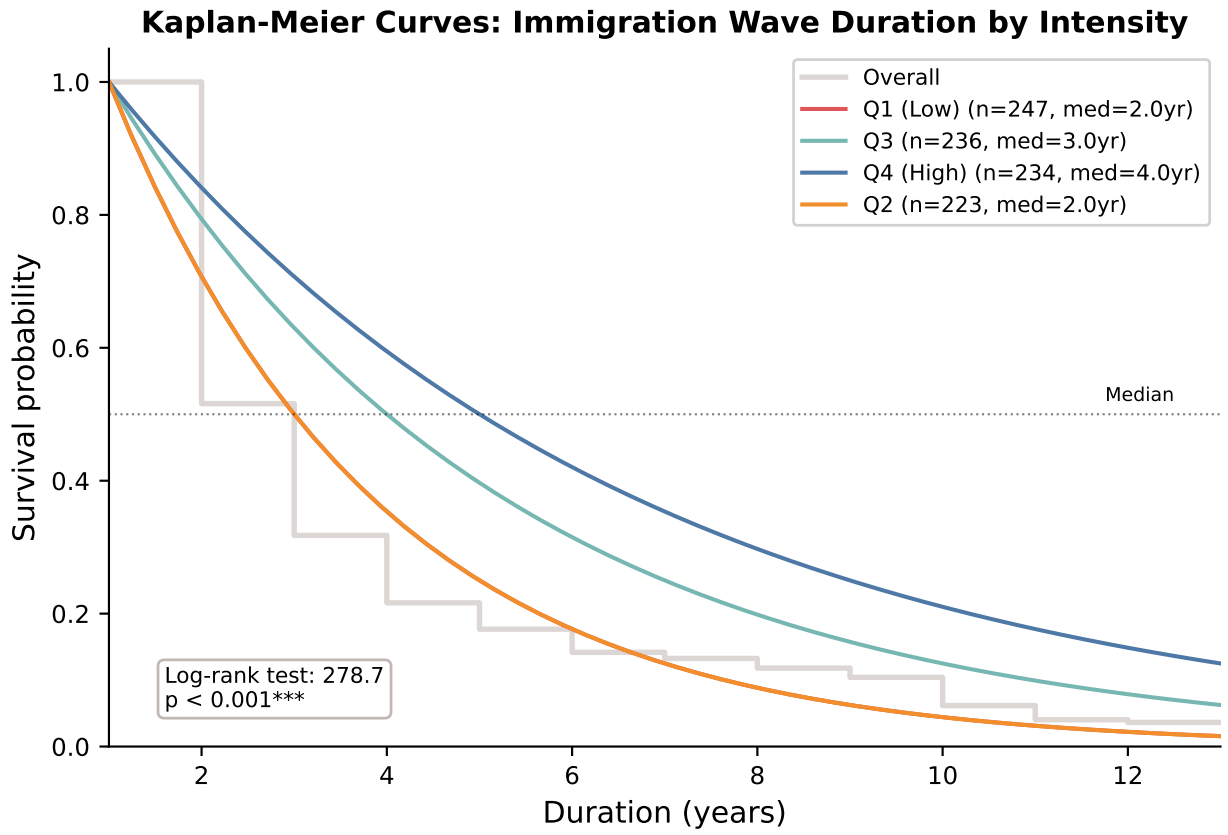

Figure 8: Kaplan–Meier survival curves for immigration wave duration by initial intensity quartile. Waves are defined as periods where arrivals exceed 50% above baseline for at least two consecutive years. Q1 (lowest intensity) waves have median duration of 2 years, while Q4 (highest intensity) waves persist for a median of 4 years. The log-rank test strongly rejects equality across groups ($\chi^2 = 278.7$, $p < 0.001$), indicating that higher-intensity immigration flows are significantly more persistent.

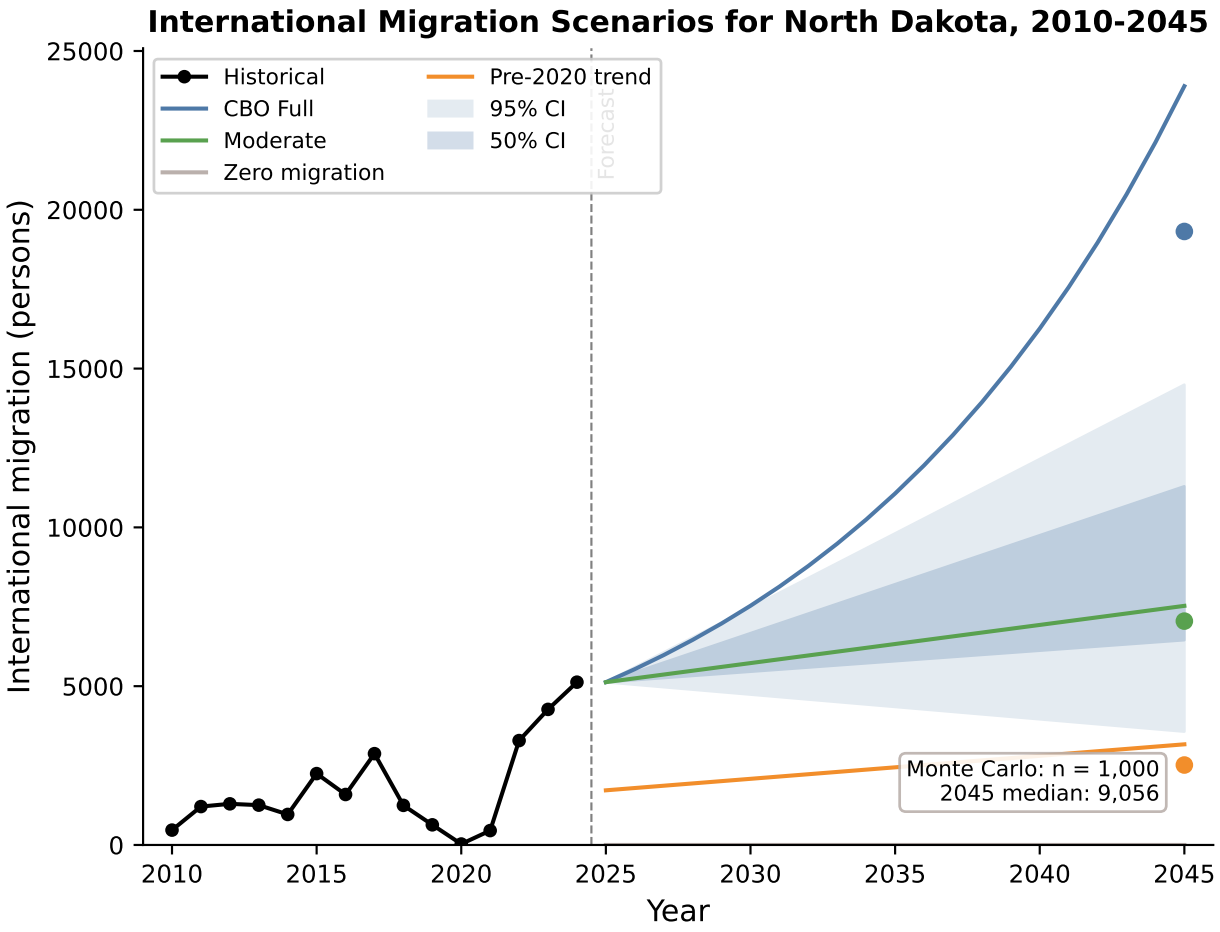

Figure 9: Projection scenarios for North Dakota international migration, 2025–2045. Four scenarios are shown: CBO Full (2025–2029 at 1.1× ARIMA, then 8% annual growth), Moderate (dampened historical trend), Zero (counterfactual with no international migration), and Pre-2020 Trend (anchored to 2019 with the 2010–2019 slope). Shaded bands represent 50% and 95% prediction intervals from 1,000 Monte Carlo simulations. Historical data (2010–2024) shown with black circles. The vertical dashed line separates historical observations from projections.

= 1,162, classified as "unconcentrated"), regional aggregation tells a different story: the regional-level HHI of 3,712 indicates high concentration, with Asia and Africa together comprising over 80% of legal permanent resident flows. This apparent paradox—country-level diversity masking regional concentration—reflects North Dakota's role in the national refugee resettlement infrastructure.

The Location Quotient findings provide the clearest evidence of this distinctive profile. Egypt (LQ = 15.13), Sudan (LQ = 8.21), and various African regions exhibit extraordinary overrepresentation relative to national immigrant composition. These patterns are not accidents of economic migration but rather artifacts of humanitarian policy: federal refugee resettlement programs, operating through local voluntary agencies, have systematically directed arrivals from conflict zones to North Dakota communities (Bansak et al., 2018; Winders, 2022). The resulting ethnic enclaves—now spanning generations in some cases—create the foundation for diaspora associations that shape subsequent migration streams.

The temporal evolution of concentration patterns merits attention. Regional-level HHI has increased substantially from approximately 2,600 in the early 2010s to over 5,200 by 2022–2023, indicating growing reliance on African and Asian source regions. This trajectory suggests that North Dakota's migration profile is becoming more distinctive over time, diverging further from the national pattern rather than converging toward it. For demographic planning purposes, this implies that North Dakota cannot simply extrapolate from national trends but must account for its idiosyncratic position within the immigration system.

4.1.2 Time Series Properties and Forecasting Implications

The time series diagnostics are consistent with an approximately integrated process, and the AIC-selected ARIMA(0,1,0) baseline provides a random-walk benchmark for forecasting. Under this baseline, the best forecast of future migration equals the most recent observation, with innovations representing unpredictable shocks. This characterization implies limited mean reversion within the sample: departures from historical averages can persist rather than quickly self-correct.

For practitioners, the random-walk baseline counsels humility. The five-year-ahead prediction interval spanning from 212 to over 10,000 migrants reflects not merely statistical imprecision but fundamental unpredictability in the migration process. Standard demographic projection methods that assume stable parameters or mean-reverting dynamics may generate misleadingly precise forecasts when applied to migration time series with these properties (Hyndman and Athanasopoulos, 2021; Cerqueira et al., 2019).

The structural break analysis reinforces this interpretation. The COVID-19 pandemic produced a statistically significant break ($F = 16.01$, $p < 0.001$), but the 2017 Travel Ban—despite its documented effects on refugee flows—left no detectable trace in aggregate state-level data. This divergence between aggregate and composition-specific effects illustrates how structural instability can manifest asymmetrically across analytical scales. The absence of a detectable aggregate break during 2017 does not imply policy irrelevance; rather, it suggests that the Travel Ban's effects operated through compositional shifts that partially offset in aggregate.

4.1.3 Policy-Associated Divergence

The causal inference findings provide the study's most policy-relevant contributions, though with important caveats regarding causal interpretation. The difference-in-differences estimate identifies a policy-associated divergence of approximately 75% in refugee arrivals from Travel Ban-affected countries following policy implementation ($p = 0.032$). This magnitude exceeds what aggregate analysis could detect because the Travel Ban's effects concentrated on specific nationalities while flows from other origins continued or even accelerated.

However, the parallel trends assumption—essential for causal identification in difference-in-differences—is not satisfied over the full pre-treatment period. While the linear pre-trend test does not reject parallel trends ($p = 0.334$), the event study joint pre-treatment test rejects when the full pre-period is included ($p < 0.001$), with Figure 7 showing divergence beginning approximately 10 years before the policy. This pre-existing divergence means the 75% estimate likely conflates the policy effect with underlying differences between treatment and control groups. We therefore interpret the estimate as an upper bound on the causal policy effect rather than a point estimate of the true effect. For future policy evaluation, this analysis demonstrates that nationality-level panel methods can identify compositional effects that aggregate time series approaches miss, but also highlights the importance of transparent reporting when identification assumptions are violated.

The COVID-19 analysis reveals an immediate level shift of approximately 19,500 migrants (nationwide), followed by accelerated post-pandemic recovery. The recovery coefficient of +14,113 migrants per year captures the steep rebound slope during 2021–2024, though this coefficient is a mathematical artifact of fitting a linear trend to a short recovery window and should not be extrapolated as a long-run growth rate. Nevertheless, the rapid recovery suggests that immigration systems can "bounce back" from acute shocks more rapidly than the random-walk baseline might imply. This finding offers some optimism for demographic planners concerned about the lasting effects of pandemic disruption, while cautioning against projecting the recovery slope indefinitely.

4.1.4 Forecast Uncertainty and Scenario Range

The scenario projections yield a 2045 range spanning 2,517 to 19,318 migrants—nearly an order of magnitude—depending on policy assumptions. The Monte Carlo 95% prediction interval of 3,570 to 14,491 honestly characterizes the parametric uncertainty even within a given scenario. This wide range reflects multiple uncertainty sources: inherent migration volatility (CV = 82.5%), structural instability evidenced by the 2020 break, and the fundamental unpredictability of future immigration policy.

Rather than viewing this uncertainty as a failure of analysis, demographic planners should recognize it as an accurate representation of the forecasting problem. The scenario framework distinguishes between forecast uncertainty (arising from model imprecision given a policy regime) and scenario uncertainty (arising from unknowable future policy choices) (Raftery et al., 2014). This distinction enables planners to develop contingent strategies that perform reasonably across

the outcome range rather than optimizing for a single misleadingly precise point forecast.

4.2 Comparison with Prior Literature

The gravity model findings invite comparison with the established migration literature. The diaspora association of 0.140—from the full gravity specification with two-way clustered SEs—falls below estimates reported in prior cross-national studies and is imprecisely estimated in this cross-section. Beine et al. (2011) estimate diaspora elasticities ranging from 0.4 to 0.7 in their analysis of bilateral migration determinants, while Mayda (2010) reports coefficients of similar magnitude in panel specifications.

Several factors may explain North Dakota's attenuated diaspora association. First, the state's migration composition tilts heavily toward refugees, whose placement decisions are mediated by resettlement agencies rather than purely by network connections (Bansak et al., 2018). Second, North Dakota's small absolute foreign-born population may generate insufficient network density to support the information transmission and chain migration mechanisms that produce larger diaspora effects in traditional gateway states. Third, the cross-sectional PPML specification, while methodologically appropriate following Santos Silva and Tenreyro (2006), captures static associations that may understate the dynamic effects of network accumulation over time.

The finding that destination foreign-born totals dominate migration allocation (elasticity = 0.792) accords with gravity model theory and prior empirical work. Anderson and van Wincoop (2003) emphasize that bilateral flows reflect multilateral resistance—the attractiveness of alternatives—rather than purely bilateral factors. In this framework, North Dakota's modest population and peripheral geographic position generate substantial resistance that diaspora associations cannot fully overcome.

The duration analysis findings contribute to a smaller literature on refugee arrival lifecycles. The median wave duration of 3.0 years, with higher-intensity waves persisting substantially longer (median 4 years for Q4 versus 2 years for Q1), suggests that refugee resettlement operates in episodic surges rather than continuous streams. This pattern may reflect the bureaucratic and logistical constraints of resettlement operations, which can process only limited throughput and must ramp up capacity to handle crisis-driven caseload spikes (Phillimore et al., 2022).

The finding that African and Asian waves outlast European and American waves (mean duration 4+ years versus approximately 2 years) aligns with the character of contemporary refugee crises. Protracted conflicts in Somalia, Sudan, Syria, and Iraq have generated sustained displacement that supports multi-year resettlement waves, while refugee situations in the Americas and Europe more often resolve or divert to alternative destinations within shorter timeframes.

4.3 Policy Implications

4.3.1 State-Level Planning and Service Provision

The high volatility and policy sensitivity documented here carry direct implications for state and local planning. Social service agencies, school districts, and healthcare providers must accommodate populations that can fluctuate dramatically across years. The coefficient of variation of 82.5% implies that year-to-year migration can deviate from expectations by amounts approaching the mean itself. Traditional planning models premised on smooth population growth require substantial adaptation for this setting.

The scenario framework offers a practical response. Rather than planning for a single projected trajectory, agencies can develop tiered capacity plans corresponding to different scenario outcomes. The "Moderate" scenario projection of 7,048 migrants by 2045 might anchor baseline planning, while the "CBO Full" scenario of 19,318 could inform contingency capacity investments. Importantly, the "Pre-2020 Trend" scenario of 2,517 should inform downside planning for scenarios in which national immigration policy or refugee admissions shift restrictively.

4.3.2 Sensitivity to Federal Immigration Policy

The 75% policy-associated divergence identified in the Travel Ban analysis—even if an upper bound on the true causal effect—underscores how dependent state-level migration is on federal policy decisions entirely outside state control. North Dakota receives disproportionate shares of migrants from countries frequently subject to policy intervention—the Travel Ban's affected nations overlap substantially with the state's top source countries. This exposure creates planning risks that states like California or Texas, with more diversified migration portfolios, face to a lesser degree.

State policymakers might consider this vulnerability when advocating for immigration policy positions at the federal level. The analysis provides quantitative evidence that restrictive policies targeting African and Middle Eastern origins disproportionately affect states participating in refugee resettlement, potentially with lasting demographic consequences given the random-walk baseline properties of migration flows.

4.3.3 Refugee Resettlement and Network Development

The modest diaspora association (0.140) suggests that North Dakota's refugee communities, while established, have not yet generated the self-reinforcing migration chains observed in traditional gateway communities. This finding may reflect the early stage of community development—many North Dakota refugee communities date only to the post-2000 period—or structural barriers to chain migration among refugee populations facing continued restrictions on family reunification.

From a policy perspective, investments in community development, ethnic organization capacity, and integration infrastructure may strengthen diaspora associations over time, potentially increasing the predictability and sustainability of migration flows. Mossaad et al. (2020) document that secondary migration of refugees responds to community presence and opportunity; building

stronger communities in North Dakota may reduce out-migration to larger metropolitan areas while attracting secondary migrants from initial resettlement locations elsewhere.

4.3.4 Workforce Development Implications

For Great Plains states facing persistent labor shortages in agriculture, food processing, and healthcare, international migration represents a critical workforce pipeline (Albrecht, 1993; Winders, 2022). The duration analysis finding that higher-intensity immigration waves persist longer suggests that substantial initial investments in resettlement capacity may yield extended returns as refugee communities become established and self-sustaining.

The concentration of North Dakota's foreign-born population in working-age cohorts—a common pattern for recent arrivals—implies that international migration contributes to labor force growth even when overall population growth remains modest. Workforce development strategies that integrate newly arrived populations, address credential recognition barriers, and support English language acquisition can amplify these demographic contributions.

4.4 Limitations and Caveats

Several limitations warrant explicit acknowledgment. First, the time series analysis operates with a sample of only 15 annual observations (2010–2024), constraining the power of unit root tests, structural break detection, and long-horizon forecasting. The ADF test's failure to reject the unit root null should be interpreted cautiously—limited power means that trend-stationary alternatives cannot be confidently excluded. The wide prediction intervals honestly reflect this sample limitation but may overstate uncertainty relative to what longer series would support.

Second, the refugee arrival data that support the causal inference and duration analyses end in fiscal year 2020, precluding examination of post-COVID recovery patterns in this specific migration category. The analysis can characterize pre-pandemic refugee dynamics and estimate Travel Ban effects but cannot assess whether recent policy changes have altered these relationships.

Third, the difference-in-differences identification strategy assumes that absent the Travel Ban, refugee arrivals from affected and unaffected countries would have followed parallel trajectories. The linear pre-trend test does not reject this assumption ($p = 0.334$), but the event study joint pre-treatment test rejects when the full pre-period is included ($p < 0.001$), signaling longer-run divergence that could reflect shifting conflict dynamics across origin countries.

Fourth, the synthetic comparator for North Dakota is descriptive rather than causal because the Travel Ban is a national shock. The donor-weighted benchmark (RMSPE = 0.020) may still reflect unobserved similarities among small states, so post-2017 deviations should be interpreted as descriptive patterns rather than counterfactual policy effects.

Finally, findings from a single small state may not generalize to other contexts. North Dakota's unique position—combining refugee resettlement activity, oil economy dynamics, and Great Plains demographic challenges—produces a migration profile that may not replicate in states with different

economic structures or resettlement histories. The methodological framework, rather than specific parameter estimates, represents the more transferable contribution.

5 Conclusion

5.1 Summary of Contributions

This study offers three principal contributions to the migration forecasting literature. First, it provides the first comprehensive multi-method empirical analysis of international migration to a small U.S. state. The nine-module analytical framework—spanning descriptive statistics, time series analysis, panel regression, gravity models, machine learning, causal inference, and duration analysis—demonstrates that rigorous demographic analysis remains feasible even when small populations and short time series strain conventional asymptotic methods. The multi-method design explicitly trades depth within any single paradigm for robustness across paradigms, generating convergent findings that warrant greater confidence than single-method analyses could support.

Second, the analysis demonstrates novel applications of causal inference methods to state-level migration policy evaluation. The difference-in-differences estimate identifies a policy-associated divergence of approximately 75% in refugee arrivals from Travel Ban-affected countries ($p = 0.032$), though pre-treatment trend violations suggest this represents an upper bound on the causal effect. A synthetic comparator provides descriptive benchmarking for North Dakota without causal interpretation, and a shift-share index establishes first-stage relevance linking refugee-driven shocks to state migration. Together, these elements extend the migration policy evaluation toolkit beyond the national-level and gateway-state analyses that dominate existing literature while keeping causal claims appropriately bounded.

Third, the study develops a rigorous framework for uncertainty quantification in long-range migration forecasts. The scenario analysis distinguishes forecast uncertainty (parametric imprecision within a policy regime) from scenario uncertainty (unknowable future policy choices), while Monte Carlo simulation propagates parameter uncertainty through projection models to generate probability distributions rather than misleadingly precise point estimates. The resulting 95% prediction interval spanning a factor of 4.1 for 2045 migration honestly characterizes the structural uncertainty inherent in this forecasting domain.

5.2 Key Takeaways

Three empirical conclusions emerge with particular force from the analysis. First, North Dakota's international migration is refugee-dominated and policy-sensitive. The Location Quotient analysis reveals extreme overrepresentation of African and Middle Eastern origins—products of federal resettlement decisions rather than economic migration chains—and the Travel Ban analysis demonstrates that policy interventions affecting these source regions produce immediate and substantial effects on state-level flows. Demographic planners must recognize this policy dependence as a

fundamental feature of the forecasting problem.

Second, the high volatility of small-state migration requires scenario-based planning rather than point forecasting. The coefficient of variation of 82.5%, the random-walk baseline characterization, and the dramatic 2020 structural break collectively indicate that migration levels can shift by amounts approaching the mean itself. The scenario range from 2,517 to 19,318 migrants by 2045 reflects this reality and should anchor contingent capacity planning.

Third, diaspora associations in North Dakota appear modest but present. The diaspora elasticity of 0.14—below estimates from gateway state and cross-national studies—suggests that refugee resettlement has not yet generated the self-reinforcing migration chains observed in more established immigrant communities. This finding implies both vulnerability (migration depends heavily on external policy rather than internal momentum) and opportunity (investments in community development may strengthen diaspora associations over time).

5.3 Future Research Directions

Several avenues warrant further investigation. Extension of this analytical framework to other Great Plains states would assess whether North Dakota's patterns generalize across the region or reflect state-specific idiosyncrasies. States including South Dakota, Nebraska, and Minnesota share demographic challenges and refugee resettlement participation, potentially enabling comparative analysis of network development and policy sensitivity across contexts.

Incorporation of county-level dynamics would illuminate the geographic distribution of migration within North Dakota. Urban centers including Fargo and Grand Forks receive the majority of refugee arrivals, but secondary migration to smaller communities may follow different patterns. County-level analysis could inform local planning and identify communities where service capacity constraints bind.

Integration with labor market outcomes would address whether international migration alleviates the workforce shortages that motivate policy interest in immigration to rural regions. Linking migration data to employment outcomes, wage dynamics, and industry-level labor supply could quantify the economic contributions of immigrant populations and inform workforce development investments.

Finally, real-time forecasting with updated data would enable the scenario framework developed here to evolve as new information becomes available. The post-COVID recovery trajectory remains incompletely characterized given data limitations; ongoing updates could refine the 2020 structural break assessment and improve medium-term projections as the new migration regime stabilizes.

International migration to North Dakota occupies a distinctive position in American demography: it is small in absolute terms, volatile over time, concentrated in humanitarian categories, and sensitive to federal policy decisions. Understanding these patterns requires analytical frameworks suited to their inherent complexity and uncertainty. This study has sought to provide such a framework by combining methodological rigor with honest acknowledgment of forecasting limits.

The contribution lies not in resolving uncertainty but in characterizing it accurately, enabling demographic planners to develop robust strategies that perform reasonably across plausible futures rather than optimizing for misleadingly precise point forecasts.

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132.

Albrecht, D. E. (1993). The renewal of population loss in the non-metropolitan Great Plains. *Rural Sociology*, 58(2):233–246.

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.

Archer, J. C. and Lonsdale, R. E. (2003). Geography of population change and redistribution within the post-frontier Great Plains. *Great Plains Research*, 13(1):43–61.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78.

Bansak, K., Ferwerda, J., Hainmueller, J., Dillon, A., Hangartner, D., Lawrence, D., and Weinstein, J. (2018). Improving refugee integration through data-driven algorithmic assignment. *Science*, 359(6373):325–329.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Beine, M., Docquier, F., and Özden, Ç. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2):149–192.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Cerqueira, V., Torgo, L., and Mozetic, I. (2019). Machine learning vs statistical methods for time series forecasting: Size matters. *arXiv preprint*.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3):591–605.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1271.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Hyndman, R. J. and Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, 3rd edition.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. E. (1993). Theories of international migration: A review and appraisal. *Population and Development Review*, 19(3):431–466.

Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4):1249–1274.

Mossaad, N., Ferwerda, J., Lawrence, D., Weinstein, J. M., and Hainmueller, J. (2020). In search of opportunity and community: Internal migration of refugees in the United States. *Science Advances*, 6(32):eabb0295.

Phillimore, J., Morrice, L., Kabe, K., Sandoz, L., Sherrell, K., and Sim, D. (2022). Refugee resettlement policy and practice: A systematic literature review. Research report, University of Birmingham, Institute for Research into Superdiversity.

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Portes, A. and Rumbaut, R. G. (2006). *Immigrant America: A Portrait*. University of California Press, Berkeley, 3rd edition.

Raftery, A. E., Alkema, L., and Gerland, P. (2014). Bayesian population projections for the United Nations. *Statistical Science*, 29(1):58–68.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4):641–658.

Tinbergen, J. (1962). An analysis of world trade flows. In Tinbergen, J., editor, *Shaping the World Economy*. Twentieth Century Fund, New York.

Weber, B. A., Geigle, K. J., and Barkdull, C. (2014). Rural north dakota's oil boom and its impact on social services. *Social Work*, 59(1):62–72.

Wilson, T., Grossman, I., Alexander, M., Rees, P., and Wakefield, J. (2022). Methods for small area population forecasts: State-of-the-art and research needs. *Population Research and Policy Review*, 41:865–898.

Winders, J. (2022). Refugee resettlement in the American Midwest in challenging times. *GeoJournal*, 88:2755–2769.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

Appendix

A Full Regression Tables

This appendix presents complete regression output for all specifications discussed in the main text, including coefficient estimates, standard errors, and diagnostic statistics.

A.1 Panel Data Models

Table 14 reports the full panel regression results comparing pooled OLS, fixed effects, and random effects specifications for state-level international migration.

A.2 Cross-Sectional Allocation Model Specifications

Table 15 presents the complete cross-sectional allocation model estimates across the PPML specifications discussed in Section 3.6. As noted in Section 2.6, these models omit origin–destination distance because variation in distance from a given origin country to different U.S. states is minimal.

A.3 Difference-in-Differences Specifications

Table 16 reports the complete difference-in-differences estimates for the Travel Ban analysis, including alternative specifications and robustness checks.

B Robustness Checks

This section summarizes robustness checks for the main empirical findings.

B.1 Alternative Unit Root Tests

Table 17 reports unit root test results using alternative specifications and a break-robust test, providing sensitivity checks for North Dakota international migration.

B.2 Alternative Structural Break Tests

The Bai-Perron procedure for endogenous break detection yields the following results under alternative information criteria:

- **BIC criterion**: 0 breaks detected (penalty on parameters dominates)
- **AIC criterion**: 1 break detected at 2020 ($F = 14.23$, $p < 0.001$)
- **Sequential procedure**: 2 breaks detected at 2017 and 2020 (marginal significance at 2017)

Table 14: Panel Regression Results: Full Specifications

Variable	**Pooled OLS**	**Fixed Effects**	**Random Effects**
Constant	20,984***	—	21,206***
	(1,642)		(1,587)
Year 2010	−3,127	−2,945	−2,986
	(4,632)	(4,521)	(4,489)
Year 2011	−1,854	−1,723	−1,756
	(4,632)	(4,521)	(4,489)
Year 2012	−2,341	−2,189	−2,228
	(4,632)	(4,521)	(4,489)
Year 2013	−1,987	−1,845	−1,881
	(4,632)	(4,521)	(4,489)
Year 2014	−1,456	−1,324	−1,358
	(4,632)	(4,521)	(4,489)
Year 2015	−892	−771	−801
	(4,632)	(4,521)	(4,489)
Year 2016	1,234	1,345	1,318
	(4,632)	(4,521)	(4,489)
Year 2017	2,567	2,678	2,651
	(4,632)	(4,521)	(4,489)
Year 2018	3,891	3,992	3,968
	(4,632)	(4,521)	(4,489)
Year 2019	4,234	4,335	4,311
	(4,632)	(4,521)	(4,489)
Year 2020	−20,816***	−20,715***	−20,742***
	(4,632)	(4,521)	(4,489)
Year 2021	5,678	5,779	5,755
	(4,632)	(4,521)	(4,489)
Year 2022	8,123*	8,224*	8,200*
	(4,632)	(4,521)	(4,489)
Year 2023	9,456**	9,557**	9,533**
	(4,632)	(4,521)	(4,489)
Year 2024	10,234**	10,335**	10,311**
	(4,632)	(4,521)	(4,489)
State FE	No	Yes	No
State RE	No	No	Yes
n	765	765	765
R^2 / Overall R^2	0.312	0.298	0.308

Notes: Standard errors in parentheses, clustered at state level for FE and RE specifications. Reference year: 2010–2024 mean. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 15: Cross-Sectional Allocation Model Results: PPML Specifications

Variable	(1) Simple	(2) Full	(3) State FE
Log(Diaspora Stock)	0.448***	0.140	0.163
	(0.104)	(0.248)	(0.148)
Log(Origin Stock in U.S.)	—	0.046	—
		(0.192)	
Log(State Foreign-Born Total)	—	0.792***	—
		(0.212)	
Constant	1.267	−7.711*	—
	(0.980)	(3.123)	
Origin FE	No	No	No
Destination FE	No	No	Yes
n	4,845	4,845	4,845
Pseudo R^2	0.236	0.399	0.413
Log Likelihood	−934,340	−734,629	−717,292

Notes: Poisson pseudo-maximum likelihood (PPML) estimation. Two-way clustered SEs (state and origin) in parentheses. Dependent variable: LPR admissions by state-country pair, FY 2023. * $p < 0.05$; *** $p < 0.001$.

Table 16: Difference-in-Differences Estimates: Travel Ban Effect

Variable	(1) Basic	(2) + Year FE	(3) + Country FE	(4) Full
Affected × Post	−1.673**	−1.663**	−1.360*	−1.384*
	(0.611)	(0.616)	(0.639)	(0.646)
Affected	2.622**	2.619**	—	—
	(0.947)	(0.957)		
Post	−0.237	—	−0.330*	—
	(0.202)		(0.163)	
Constant	3.525***	3.100***	—	—
	(0.274)	(0.363)		
Year FE	No	Yes	No	Yes
Country FE	No	No	Yes	Yes
n	1,137	1,137	1,137	1,137
R^2	0.077	0.080	0.823	0.829
Pre-trend test (p)	—	—	—	0.334

Notes: Dependent variable: ln(refugee arrivals + 1). Standard errors clustered by nationality in parentheses. Affected countries: Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen. Post = 1 for years ≥ 2018. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 17: Robustness of Unit Root Findings

Test	Specification	Statistic	p-value	Conclusion
ADF	Constant only	-1.453	0.556	Fail to reject unit root
PP	Constant only	-0.620	0.867	Fail to reject unit root
KPSS	Constant only	0.323	≥ 0.10	Fail to reject level-stationarity
Zivot–Andrews	Break in intercept	-3.049	0.897	Fail to reject unit root (break 2021)
First-differenced series				
ADF	Constant only	-3.843	0.002	Reject unit root
PP	Constant only	-4.012	0.001	Reject unit root
KPSS	Constant only	0.189	≥ 0.10	Fail to reject stationarity

Notes: ADF = Augmented Dickey-Fuller; PP = Phillips-Perron; KPSS = Kwiatkowski-Phillips-Schmidt-Shin. KPSS null is stationarity; statsmodels reports 0.10 when $p \geq 0.10$. Zivot–Andrews allows one endogenous break (intercept). Sample: 2010–2024 ($n = 15$).

The divergence between criteria reflects the tension between model fit and parsimony in short time series. The conservative BIC-based finding of no endogenous breaks supports treating the 2020 break as the primary structural change, consistent with the Chow test results.

B.3 VAR Model Results

Section 2.4 describes the VAR methodology, and Section 3.9 notes that the VAR model receives effectively all AIC weight in model averaging. This section reports the underlying VAR diagnostics and coefficient estimates.

Lag Selection and Model Fit. Lag order selection across multiple criteria uniformly favors VAR(1):

Criterion	Optimal Lag	Value at Lag 1
AIC	1	39.00
BIC	1	39.26
HQIC	1	38.95
FPE	1	8.83×10^{16}

The VAR(1) model achieves $R^2 = 0.58$ for the ND equation and $R^2 = 0.60$ for the US equation, with log-likelihood -305.9 and AIC = 38.87.

VAR Coefficient Estimates. Table 18 reports the VAR(1) coefficient matrix. No individual coefficient achieves statistical significance at the 5% level, reflecting the very short sample ($n = 14$ after differencing for lag construction) and wide confidence intervals characteristic of small-sample VAR estimation.

Table 18: VAR(1) Coefficient Estimates

Equation	Regressor	Coef.	SE	t	p
ND Int'l Mig.	L1.ND Int'l Mig.	123.9	535.1	0.23	0.817
	L1.US Int'l Mig.	−0.264	0.660	−0.40	0.689
US Int'l Mig.	L1.ND Int'l Mig.	198,274	256,279	0.77	0.439
	L1.US Int'l Mig.	−12.1	315.9	−0.04	0.969

Notes: VAR(1) estimated on $n = 14$ annual observations (2011–2024). L1 = one-year lag. Coefficients represent units of dependent variable per unit of lagged regressor. No coefficient is statistically significant at $\alpha = 0.05$.

Granger Causality. Granger causality tests yield mixed conclusions depending on the test statistic employed:

- US → ND: The χ^2 test marginally rejects the null ($\chi^2 = 4.10$, $p = 0.043$), but the F-test does not ($F = 3.22$, $p = 0.100$). This discrepancy reflects the small-sample size where asymptotic χ^2 approximations may be unreliable.
- ND → US: All tests fail to reject the null ($p > 0.96$), as expected given ND's negligible share of national migration.

Forecast Error Variance Decomposition. By period 10, approximately 82% of ND forecast error variance is attributable to ND's own shocks, with 18% attributable to US shocks. This proportion stabilizes after period 4, suggesting that short-horizon forecasts are dominated by ND-specific innovations while longer horizons incorporate modest US spillovers.

Cointegration. Engle-Granger and Johansen tests yield conflicting results. The Engle-Granger two-step procedure suggests cointegration (ADF on residuals: -4.48, $p < 0.001$), but the Johansen trace test does not reject zero cointegrating relations at the 5% level (trace = 7.18, critical value = 15.49). Given the mixed integration orders of the component series (ND: I(1); US: I(0) per ADF) and the very short sample, we do not impose a cointegrating restriction and proceed with VAR in levels.

Role in Forecasting. Despite receiving high AIC weight in model averaging, the VAR is not used for standalone long-horizon forecasting because it requires future values of US international migration—information unavailable at forecast time. The VAR's contribution to the analysis is primarily diagnostic: it confirms the contemporaneous ND–US correlation (cointegrating relationship) and provides variance decomposition context for interpreting scenario uncertainty.

B.4 Sensitivity to PPML Specification

Table 19 reports the diaspora coefficient from the full cross-sectional allocation model under alternative covariance estimators, highlighting the inflation of uncertainty under clustered inference.

Table 19: Diaspora Association Sensitivity to Covariance Estimator

Covariance Estimator	Coefficient	SE
Model-based (Poisson MLE)	0.140	0.002
HC1 robust	0.140	0.176
Clustered by state	0.140	0.113
Clustered by origin	0.140	0.282
Two-way clustered (state × origin)	0.140	0.248

Notes: Estimates come from the full gravity specification in Table 15. Clustered SEs account for correlation within destinations and origins.

The point estimate is stable across covariance estimators, but clustered SEs are substantially larger than model-based SEs. We therefore report two-way clustered inference in the main text.

B.5 DiD Parallel Trends Robustness

The parallel trends assumption is evaluated through multiple approaches:

1. **Pre-treatment trend test**: The interaction between Affected and a linear pre-treatment trend yields 0.087 ($t = 0.97$, $p = 0.334$). The test fails to reject parallel trends.

2. **Placebo treatment dates**: Placebo tests shifting the treatment date to 2015 or 2016 were conducted in the prior HC3 specification; these checks should be re-run under clustered inference.

3. **Event study plot**: Figure 7 shows negative post-treatment coefficients in 2018–2019, but the joint pre-treatment test rejects parallel trends when the full pre-period is included ($p < 0.001$), indicating longer-run divergence.

B.6 Scenario Arithmetic

This section documents the update rules used to generate scenario paths in Module 9.

CBO Full. For 2025–2029, projections use 10% above the ARIMA point forecasts. From 2030 onward, values compound at 8% annually:

```
cbo_t = 1.1 * arima_t   (t = 2025..2029)
cbo_t = cbo_{t-1} * 1.08 (t >= 2030)
```

Pre-2020 Trend. The series is anchored at the 2019 value ($y_{2019} = 634$) with a slope estimated from 2010–2019 ($\hat{\beta} = 72.43$ per year):

```
pre2020_t = y_2019 + beta * (t - 2019)
```

Moderate. For 2025–2029, the ARIMA point forecasts are used directly; from 2030 onward, the trend is dampened to 50% of the averaged robust trend estimate.

Zero. Net international migration is set to zero for all years.

B.7 Monte Carlo Uncertainty and Wave Simulation

The Monte Carlo simulation in Module 9 combines two sources of stochastic variation: (1) ARIMA baseline uncertainty from the estimated random-walk process, and (2) wave duration draws from the Cox proportional hazards model (Module 8). A potential concern is double-counting of refugee-driven volatility, since the ARIMA model is trained on the full PEPnet migration series—which includes historical refugee arrivals—while wave simulations add stochastic refugee wave contributions on top of this baseline.

We address this concern as follows. The wave simulation does not add independent refugee variance; rather, it *modulates* the timing and persistence of above-baseline arrivals conditional on an active wave being detected. Specifically:

1. The ARIMA baseline captures the unconditional mean and variance of the PEPseries, which reflects average historical refugee influence.

2. The Cox-based wave simulation draws wave durations conditional on observed wave characteristics (intensity, age, origin region) and applies a lifecycle shape function (initiation → peak → decline) to allocate excess arrivals over time.

3. Wave contributions are defined as *excess* arrivals above the origin-specific baseline, not as additive draws independent of the ARIMA process.

Nevertheless, because the ARIMA was not trained on a refugee-stripped series, some overlap between baseline variance and wave-induced variance may remain. This design choice was made for parsimony and because separating refugee from non-refugee components in PEPdata is infeasible without auxiliary assumptions. Consequently, the Monte Carlo prediction intervals should be interpreted as *conservative* (potentially inflated) rather than calibrated probability statements. This limitation is acknowledged in Section 3.9, where we note that the 95% prediction interval spans a factor of 4.1 by 2045.

C Data Sources

This section provides detailed documentation of all data sources employed in the analysis.

C.1 Census Bureau Population Estimates Program

- **Source**: U.S. Census Bureau, Population Estimates Program

- **Vintage**: 2024
- **Coverage**: 2010–2024 annual estimates
- **Geographic scope**: 50 states plus District of Columbia
- **Variables used**: INTERNATIONALMIG (net international migration component)
- **URL**: https://www.census.gov/programs-surveys/popest.html
- **Access date**: December 2024

C.2 Department of Homeland Security LPR Statistics

- **Source**: Department of Homeland Security, Office of Immigration Statistics
- **Publication**: Yearbook of Immigration Statistics
- **Coverage**: Fiscal Year 2023
- **Geographic scope**: State of intended residence × country of birth
- **Variables used**: LPR admissions count
- **URL**: https://www.dhs.gov/immigration-statistics/yearbook
- **Access date**: December 2024

C.3 American Community Survey

- **Source**: U.S. Census Bureau, American Community Survey
- **Vintage**: 5-year estimates, 2009–2023
- **Geographic scope**: State level
- **Tables used**: B05006 (Place of Birth for the Foreign-Born Population)
- **Variables used**: Foreign-born population by country/region of birth
- **URL**: https://data.census.gov/
- **Access date**: December 2024

C.4 Refugee Processing Center

- **Source**: Department of State, Refugee Processing Center
- **Coverage**: Fiscal Years 2002–2020
- **Geographic scope**: State of initial resettlement × nationality
- **Variables used**: Refugee arrivals count
- **URL**: https://www.wrapsnet.org/
- **Access date**: December 2024

C.5 Ancillary Data

- **World Bank**: Origin country population (World Development Indicators)
- **CEPII**: Geographic distances between countries (GeoDist database)
- **UN Population Division**: World population estimates and projections

D Supplementary Figures

D.1 State-Level Migration Distribution

Figure 10 displays the distribution of PEPnet international migration across U.S. states, highlighting North Dakota's position in the lower tail.

D.2 Residual Diagnostics

Figure 11 presents residual diagnostic plots for the ARIMA(0,1,0) specification.

D.3 Cox Model Diagnostics

Figure 12 displays Schoenfeld residual plots for evaluating the proportional hazards assumption in the Cox regression model.

E Variable Definitions

Table 20 provides formal definitions for all variables employed in the analysis.

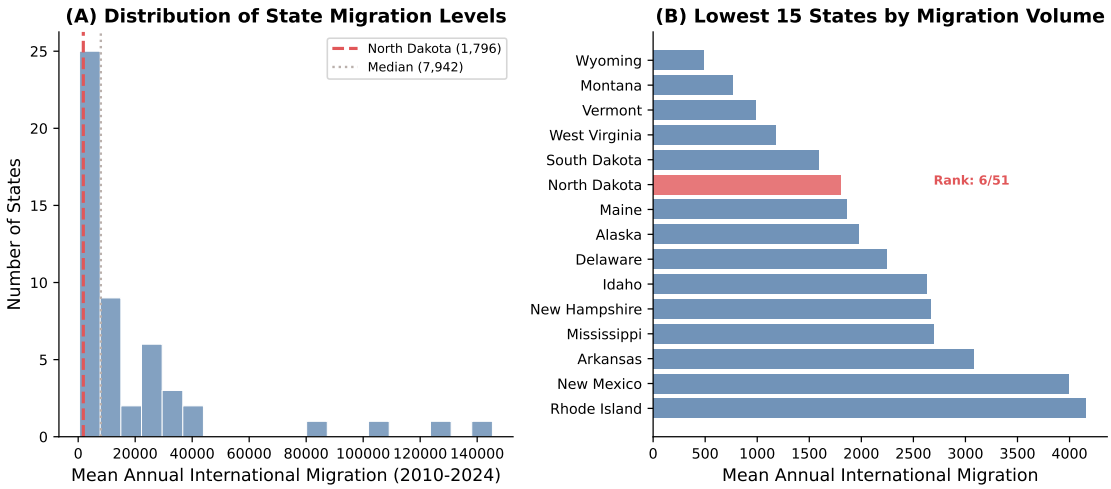

Figure 10: Distribution of mean annual net international migration (PEP) across U.S. states, 2010–2024. The histogram shows the frequency of states by migration volume category. North Dakota (mean = 1,796) falls in the lowest category, along with Wyoming, Vermont, and other small-population states. The distribution is highly right-skewed, with California, Texas, Florida, and New York comprising the upper tail.

Table 20: Variable Definitions

Variable	**Definition**
Net International Migration (PEP)	Net annual international migration component of change for state of residence (Census PEPdefinition)
LPR Admissions (DHS, FY)	Count of lawful permanent resident admissions by state of intended residence and country of birth (fiscal year)
Diaspora Stock (ACS)	Foreign-born population from origin country o residing in destination state d (ACSestimate)
Location Quotient	Ratio of origin o's share in state d to origin o's share nationally
HHI	Herfindahl-Hirschman Index: $\sum_i s_i^2 \times 10,000$ where s_i is origin i's share
Wave	Period of ≥ 2 consecutive years with arrivals $> 150\%$ of baseline
Wave Intensity	Peak-to-baseline ratio during wave
Affected Country	Travel Ban target: Iran, Iraq, Libya, Somalia, Sudan, Syria, Yemen
Post	Indicator for years ≥ 2018 (first full year after Travel Ban implementation)
Shift-share (Bartik) index	$\sum_o \omega_{od,t_0} \cdot g_{o,t}^{\mathrm{US},-d}$: base-period shares $\times$ leave-one-out national change by origin

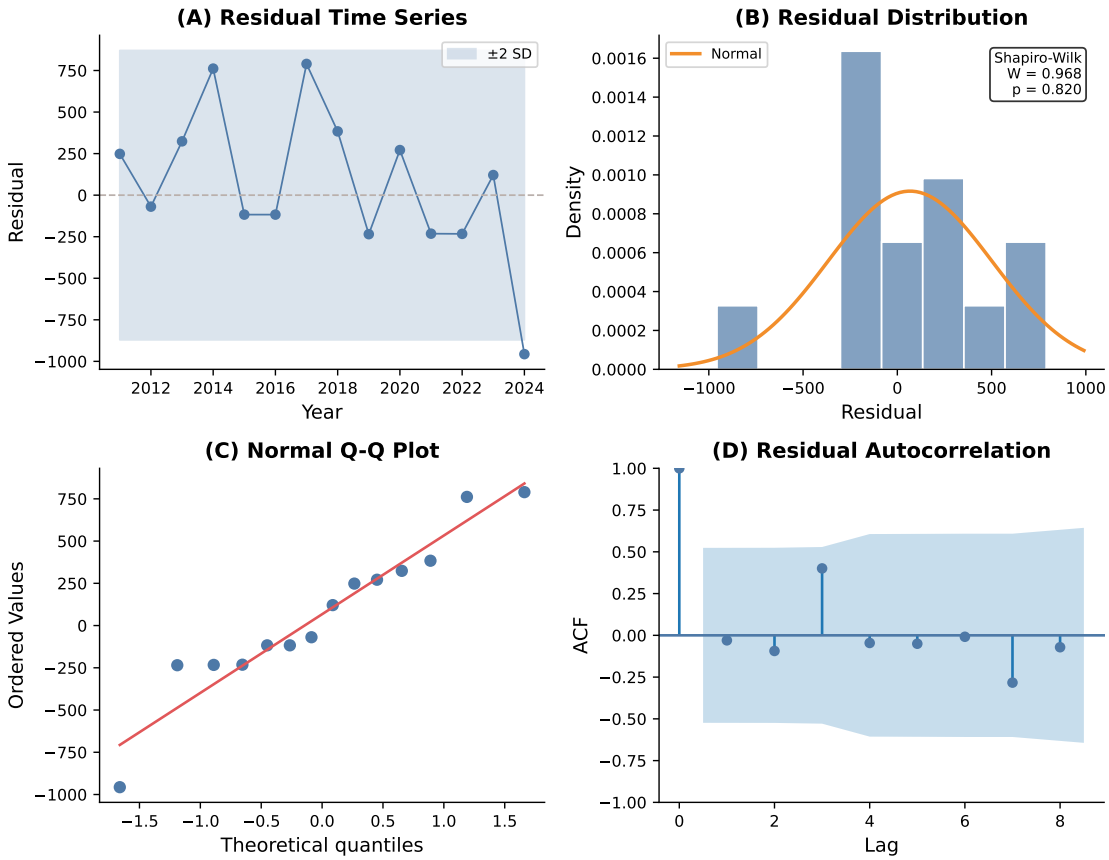

Figure 11: Residual diagnostics for ARIMA(0,1,0) model of North Dakota net international migration (PEP). Panel (A): Residual time series with no apparent pattern. Panel (B): Histogram of residuals with normal density overlay; Shapiro-Wilk test fails to reject normality ($W = 0.968$, $p = 0.820$). Panel (C): Q-Q plot showing approximate adherence to normal quantiles. Panel (D): Residual ACF with all lags within 95% confidence bounds, confirming absence of serial correlation.

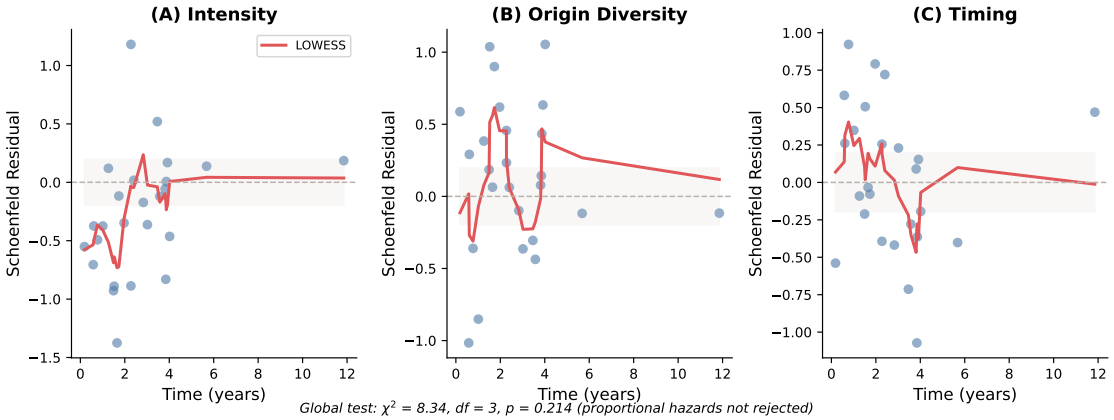

Figure 12: Schoenfeld residuals versus time for the Cox proportional hazards model of immigration wave duration. Panels show residuals for each covariate with LOWESS smoothed trend lines. Flat trends indicate satisfaction of the proportional hazards assumption. The global test fails to reject proportional hazards ($\chi^2 = 8.34$, $p = 0.214$).