Migration and Fertility Decisions over the Life Course*

Jaehyun Choi† Coady Wing‡

July 10, 2026

1 Datasets

List the datasets that you will request in SAP before starting the application.

- American Community Survey (ACS) (RDG 691) 2000–2024
- BOC PIK Crosswalk American Community Survey (ACS) (RDG 2595) 2001–2024
- Master Address File extract (MAFX) 2008–2026
- Master Address Administrative Reference file (MARF) 2000–2023
  - BOC Master Address File Auxiliary Reference File (MAFARF) (RDG 720) 2000–2023
  - BOC Master Address File–Multiple MAFIDs (RDG 18744) 2013–2021, 2023–2025
- Residence Candidates File (RCF) 1999–2023
  - Longitudinal Employer-Household Dynamics (LEHD) Person-level Files Residence Components (RDG 7438) – including all states
- Survey of Income and Program Participation (SIPP) (RDG 620) 1984–2024 (all panels)
- BOC Consolidated PIK Crosswalk SIPP 1984–2025
- Current Population Survey Annual Social and Economic Supplement (CPS ASEC) (RDG 603) 1973–2025

*The authors declare no known conflicts of interest, financial or otherwise, related to this work.

†O'Neill School of Public and Environmental Affairs, Indiana University - Bloomington. E-mail: jc234@iu.edu
‡O'Neill School of Public and Environmental Affairs, Indiana University - Bloomington. E-mail: cwing@iu.edu

- BOC PIK Crosswalk Current Population Survey-ASEC (RDG 2666) 1973–2025
- CPS Fertility Supplement (biennial, even years) (RDG 731) 1998–2022
- BOC Consolidated PIK Crosswalk CPS (RDG 2593) 1973–2025
- Census Household Composition Key File 2016–2023
- Decennial Census 2000 and 1990 (RDG 690)
- BOC PIK Crosswalk Census Unedited File (CUF) 2000 (RDG 7908)
- EIF Demographic Spine (2025 release)
- EIF Residential Information 1999–2024
- Longitudinal Business Database (LBD) 1976–2023

2 Research Description

2.1 Project Title: Migration and Fertility Decisions

2.2 Project Duration: Five years

2.3 Funding

Indiana University.

2.4 Timeline

- Activity A: Identify women ages 15–45 in ACS and construct the analytic sample.
- Activity B: Link ACS records to residence files using PIKs; merge to MARF, MAFX, and RCF to construct longitudinal residence histories, and link to EIF Residential Information as an additional residence-history and environmental-exposure source.
- Activity C: Assess the quality of administrative residence assignments relative to ACS-recorded addresses using MARF, RCF, and EIF; document discrepancies and separated-parent cases in a technical memorandum (Criterion 5).

- Activity D: Construct retrospective fertility outcomes and destination fertility environment measures; identify migration events from residence histories.
- Activity E: Estimate retrospective mover comparison models and coarse post-move timing profiles (Criterion 11).
- Activity F: Conduct robustness checks, including LBD shock-subsample analyses, and subgroup analyses.
- Activity G: Prepare draft manuscript; submit outputs for Census Bureau disclosure review.
- Activity H: Finalize and submit research paper; deposit technical memoranda on Census Bureau server.

Table 1: Project Timeline (60 Months)

Task	**Y1**	**Y2**	**Y3**	**Y4**	**Y5**
Activity A: Construct ACS analytic sample	X				
Activity B: Link ACS to MARF, MAFX, RCF, EIF	X				
Activity C: Data quality assessment (C5)	X	X			
Activity D: Construct outcomes and migration events		X			
Activity E: Estimate retrospective mover comparison models (C11)		X	X		
Activity F: Robustness checks, including LBD shock-subsample analyses, and subgroup analyses			X	X	
Activity G: Disclosure review and manuscript preparation				X	
Activity H: Submit paper; deposit technical memos (C5, C11)				X	X
Submit annual progress report (Years 1-4)	X	X	X	X	
Submit final progress report					X
Submit post-project certification					X

3 Research Question: [/3,800 characters]

Observed fertility rates vary substantially across places within the United States. In 2024, the District of Columbia reported 39.8 births per 1,000 women ages 15–44, compared with 66.7 births per 1,000 women ages 15–44 in South Dakota, a difference of 26.9 births per 1,000 women and a rate roughly 68 percent higher in South Dakota (March of Dimes, 2026a,b). The reasons for spatial differences in fertility rates are not well understood. It is possible that geographic areas have *causal effects* on fertility decisions. Causal place effects on fertility could reflect local amenities or social norms that affect the costs and benefits of having children. Alternatively, the variation in fertility rates across the country might reflect selection bias that arises because people with preferences for larger (or smaller) family sizes migrate into certain parts of the country.

The study we propose will use multiple linked data sources to study the connection between two important demographic processes at the individual level: (i) internal migration, and (ii) fertility and family formation. The project will help improve the data infrastructure needed to measure individual migration histories and to connect these histories with information on fertility and family outcomes. We will exploit within-person migration histories to help identify the causal effects of place on fertility using a version of the *mover designs* that have been used in previous research to measure place effects on mortality, health care spending, and intergenerational economic mobility (Finkelstein et al., 2016, 2021; Chetty and Hendren, 2018a,b).

1. How accurately do alternative administrative residence records (including MARF, RCF, and EIF) measure the place of residence for women near the time of childbirth events, and does the quality of these residential measurements vary systematically by household structure, geographic area, and socioeconomic status?

2. What is the causal effect of migrating from a low-fertility area to a high-fertility area?

3. Are place effects on fertility associated with geographic variation in fertility and family-related amenities or with markers of social norms related to family size?

Answering these questions requires restricted-use Census Bureau microdata. In the public-use ACS files, geographic identifiers are coarsened to Public Use Microdata Areas (PUMAs) and

individuals cannot be followed across survey rounds. The proposed analysis links restricted ACS records to administrative residence files (MARF, MAFX, and RCF) using Protected Identification Keys (PIKs) that have already been created by the Census Bureau. Linking the ACS data with these residential files will allow us to construct the longitudinal residence histories for individuals in the ACS.

4 Demonstrated need: [/10,000 characters]

This project requires access to restricted-use Census Bureau microdata to construct longitudinal residence histories, reconstruct retrospective fertility measures from household roster data, and evaluate the quality of administrative residence assignments around childbearing events. The research questions cannot be addressed using public-use data for three primary reasons.

First, public-use versions of major Census surveys suppress or substantially coarsen geographic identifiers and do not support longitudinal linkage at the individual level. The central research questions require tracking individuals across locations over time and identifying both the timing and destination of residential moves. Public-use ACS and Decennial Census files do not provide sufficiently detailed geography to measure residential mobility, nor do they allow reliable identification of repeated moves by the same individual across years. Person identifiers in public-use surveys are designed to be unique within survey year and cannot be used to follow individuals longitudinally. This project therefore requires restricted-use person identifiers to link individuals across survey waves and to administrative residence records in order to construct consistent longitudinal residence histories. Such residence histories cannot be constructed using public-use files.

Second, accurate measurement of fertility responses to migration requires precise observation of household composition and child ages at the interview date, combined with dated residential moves from administrative records. The project's fertility measures are reconstructed retrospectively from the ACS household roster using co-resident children linked to the woman through the roster and their ages in completed years. Public-use data do not permit the person-level linkage needed to combine these household data with MARF- and RCF-based move histories, and they often suppress or coarsen family-structure variables in ways that reduce precision. Restricted-use data allow more accurate identification of co-residence with children, household relationships, and the age structure

of household members. This level of detail is necessary to assign births to pre-move and post-move windows and to distinguish true fertility responses from changes in household reporting.

Third, the proposed analysis would fail with public-use data for reasons that go beyond sample size alone. Most critically, the mover-based identification strategy requires precise measurement of destination fertility environments at the commuting-zone level. Public-use ACS files identify respondents' geography only at the PUMA level. PUMAs do not correspond to commuting zones or other economically meaningful sub-state geographic units and cannot be reliably mapped to them. Without commuting-zone geography, it is not possible to assign accurate fertility environment measures to movers' destinations, and the key source of identifying variation in the empirical strategy is unavailable. As noted above, constructing the residence histories additionally requires person-level identifiers (PIKs) to link ACS records to the MARF, MAFX, and RCF — identifiers that are restricted to the RDC environment and absent from public-use files. The Criterion 5 component of the project likewise requires restricted-use linkage, because evaluating whether MARF, RCF, and EIF residence assignments are consistent with ACS-recorded addresses is only possible when the same individual can be matched across survey and administrative files. Beyond these fundamental constraints, restricted-use data also provide substantially greater statistical power. The proposed analysis focuses on a relatively narrow population — women of childbearing age who experience a residential move across commuting zones — and requires sufficient observations to estimate heterogeneous effects by age at move, prior parity, destination type, and move distance. The larger samples available in the restricted environment are essential for the precision these subgroup analyses require.

In addition to these three reasons, restricted-use establishment-level data are required for a supplementary employer-side shock analysis described in the Methodology. The Longitudinal Business Database (LBD) is used to identify large, discrete establishment births and expansions and to distinguish these events from routine aggregate employment changes; public-use business tabulations do not support this level of detail. These LBD-derived measures are aggregated to commuting-zone-by-industry-by-year cells to classify local labor-demand shocks and are not linked to individual movers or to specific firms in the analytic fertility file.

5 Study Population: [/10,000 characters]

The study population consists of women of childbearing age observed in restricted-use ACS records who can be linked via person identifiers (PIKs) to administrative residence histories. The primary analytic population includes women approximately ages 15 to 45 who experience at least one residential move within the United States and are observed in the ACS after that move.

The unit of analysis is the mover. Each woman contributes one ACS interview, from which fertility is measured retrospectively using the household roster, the ages of co-resident children, and the observed move year from the administrative residence history. These data are used to construct pre-move parity and post-move childbearing within a fixed exposure window. Individuals are included in the analytic sample if their residence history can be constructed, if key demographic characteristics are observed, and if the ACS interview occurs sufficiently long after the move to satisfy the common post-move window restriction. Individuals with incomplete linkage or insufficient residence history are excluded.

The analysis focuses on internal migrants, with identification relying on comparisons among movers rather than between movers and non-movers. Specifically, the study compares women originating from the same locations and with similar pre-move fertility histories who relocate to destinations with systematically different fertility environments. Non-movers may be included in descriptive analyses to characterize baseline fertility patterns, but the primary identifying variation comes from cross-destination differences among movers. Non-movers are also used to construct destination fertility environment measures but do not contribute to identifying variation in the main specifications.

Residential moves are defined using administrative address histories derived from Census address infrastructure. The study population is drawn from repeated cross sections of the ACS that are linked to administrative residence records, rather than from a survey panel that follows the same woman over time. This structure allows retrospective fertility measures to be assigned relative to migration timing from a single ACS interview while conditioning on origin location, pre-move fertility, and demographic characteristics.

6 Project Abstract: [/3,800 characters]

Fertility decisions are shaped by both individual economic incentives and the social environments in which people live. A growing literature suggests that local social norms, community context, and place-based institutional factors exert causal effects on childbearing, but identifying these effects requires longitudinal data that are unavailable in public-use surveys.

This project studies how residential mobility affects fertility and family size decisions over the life course in the United States. The analysis focuses on women of childbearing age and estimates the effect of destination location on subsequent fertility by comparing movers to destinations with systematically different fertility environments, measured using age-specific fertility rates at the commuting-zone level. The sample is constructed by linking restricted-use American Community Survey (ACS) microdata — which provide the household roster, household composition, and demographic characteristics needed to reconstruct retrospective fertility measures — to three administrative residence files: the Master Address File extract (MAFX), the Master Address Administrative Reference file (MARF), and the Residence Candidates File (RCF). These files track individuals' residential addresses over time via person identifiers (PIKs) and address identifiers (MAFIDs), enabling construction of longitudinal migration histories that cannot be produced from public-use data alone.

By focusing on internal migration rather than international migration, the project isolates changes in place-based exposure within a common national institutional framework. The empirical strategy exploits variation in the timing and destination of residential moves while conditioning on origin location, pre-move fertility, and observed demographic characteristics. The central hypothesis is that relocation to destinations with higher fertility environments increases subsequent childbearing, and that these responses reflect the causal influence of local norms and institutional contexts. The analysis also examines heterogeneity by age at move, prior parity, and move characteristics such as distance.

The project contributes to the literature on place-based effects and family formation by extending mover-based identification strategies, widely used in health and mortality research, to fertility outcomes. A complementary contribution evaluates the quality of administrative residence records around the time of childbearing. If the MARF or RCF systematically misassign women's

locations—for example, in cases involving children of separated parents who may be associated with multiple addresses—the fertility environment measure may be mismeasured for precisely the women whose behavior is under study, biasing the main estimates. The analysis compares administrative residence assignments with ACS-recorded addresses by fertility status and household composition, producing internal technical documentation for Census Bureau use (Criterion 5). The main causal estimates will be submitted for publication in a top general-interest economics journal (Criterion 11).

7 Time, Geographic, and Other Units Requested [/10,000 characters]

We are requesting access to years 2000–2024 for the ACS-linked core analysis. Earlier survey years are requested for the SIPP and CPS ASEC validation files, which extend back to 1984 and 1973, respectively. The LBD is requested for 1976–2023 to construct local labor-demand shock measures used in the employer-side shock robustness analysis. These establishment-level records are aggregated to commuting-zone-by-industry-by-year cells before being merged to the mover sample.

Molloy et al. (2011) document that approximately 1.5 percent of the U.S. population moves between Census regions each year from 1980 to 2010. Also, about 5 to 6 percent of the population moves across a county boundary annually, a distance that is often sufficient to entail meaningful changes in local housing and labor market environments. These migration flows are substantial relative to other forms of economic adjustment and persist even as aggregate migration rates decline over time. The requested ACS years (2000–2024, or the most recent year available at the time of analysis) span more than two decades of residential mobility and fertility behavior in the United States, providing a sufficiently long observation window to track women through multiple life-course stages and to estimate how changes in place of residence affect subsequent childbearing decisions. The earlier SIPP and CPS ASEC years are requested only for validation and measurement-calibration analyses, while the historical LBD years are requested only to construct local employment-shock measures and pre-period industry-location baselines.

The ACS, SIPP, and CPS ASEC (and the BOC PIK Crosswalk CPS-ASEC) are requested for

the full year range available at RDC because these surveys supply demographic detail — race and ethnicity, household composition, household relationships, age structure of co-resident children, marital status, and education — that is not available in the administrative residence and address files used elsewhere in the project. Requesting all available vintages additionally enlarges the sample available for fertility-relevant subgroup analyses (e.g., by age at move, prior parity, race and ethnicity, and move type), which is important for satisfying Census Bureau disclosure thresholds when reporting these subgroup estimates.

The project requires sub-state geographic detail, including identifiers at the level of commuting zones or comparable aggregations, in order to identify residential moves and characterize meaningful changes in place-based exposure. Fine geographic detail is essential to distinguish moves that plausibly alter exposure to local fertility environments from within-area mobility that does not. Public-use geographic identifiers are insufficient for this purpose because they mask cross-area migration and do not allow consistent construction of longitudinal residence histories.

The unit of observation is the individual woman, with fertility outcomes measured using household-level information on the presence and timing of children. Residence histories are constructed at the address level using administrative records and aggregated to relevant geographic units for analysis.

8 Methodology

8.1 Overview

The economics of fertility emphasizes how costs and incentives shape childbearing decisions across the life course (Becker and Lewis, 1973; Doepke et al., 2023). A substantial empirical literature documents that fertility responds to financial incentives, parental leave policies, and labor market conditions (Milligan, 2005; Lalive and Zweimüller, 2009; Schaller, 2016). However, fertility rates vary substantially across US communities, and aggregate birth rates have declined persistently since the Great Recession in ways that economic incentives alone do not fully explain (Kearney et al., 2022). A complementary channel operates through local social norms and peer behavior: quasi-experimental evidence shows that media exposure and workplace peer effects can independently shift individual childbearing decisions (Ferrara et al., 2012; Kearney and Levine, 2015; Ciliberto et al., 2016). This project focuses on this second channel, using residential mobility to identify how

changes in local fertility environments affect subsequent childbearing.

The empirical strategy adopts a mover-based design to estimate how changes in place of residence affect fertility and family size decisions over the life course. The core idea is to exploit variation in destination environments among individuals who relocate, comparing fertility outcomes for women with similar origins and pre-move characteristics who move to destinations with systematically different fertility environments.

The analysis proceeds in three main stages. First, the project constructs longitudinal residence histories for individuals by linking demographic survey records to administrative address-based residence information available within the Census Bureau's data infrastructure. These residence histories are used to identify the timing, origin, and destination of residential moves and to characterize changes in place-based exposure over time. This step establishes the migration events that define changes in local context and allows classification of moves by distance and type.

Second, residence histories are linked to ACS survey records to reconstruct retrospective measures of childbearing timing and parity from the ages and household relationships of co-resident children linked to the woman through the household roster, combined with the observed move year T_i from the administrative residence panel. This yields measures of pre-move parity and post-move childbearing within a common observation window for each mover from a single ACS interview. Because the ACS does not collect a children-ever-born question after 1998, this reconstruction relies on household roster structure rather than a direct fertility-history question.

Third, the analysis estimates retrospective mover comparisons that relate post-move childbearing within a fixed window to changes in fertility environments between origin and destination, conditioning on pre-move parity, origin location, age, and other demographics. Identification comes from comparing women who originate from similar locations and have similar pre-move fertility histories but relocate to destinations with systematically different fertility environments. This approach parallels mover-based designs used to estimate place effects in health and mortality outcomes (Finkelstein et al., 2016, 2021).

The analysis further examines heterogeneity in fertility responses by age at move, prior parity, and characteristics of moves, such as distance and housing-related changes. Robustness checks assess sensitivity to alternative definitions of residence histories, migration events, and fertility measures.

8.2 Sample Construction

The proposed research links a core set of restricted-use Census Bureau data products centered on the 2000–2024 ACS vintages used for the main mover design and the 2000–2023 administrative residence panel, and supplements them with validation files where available. The residence panel is constructed from three administrative residence files. The Master Address File extract (MAFX) is the Census Bureau's annual inventory of residential addresses, with each address assigned a unique identifier (MAFID). The Master Address Administrative Reference file (MARF)[1] links individuals to addresses over time by recording combinations of person identifiers (PIKs), MAFIDs, and calendar years, enabling construction of longitudinal residence histories. The Residence Candidates File (RCF) provides a harmonized annual residence assignment for each person: for 1999–2010, residential geography is derived from the Composite Person Record (CPR); for 2012 and later, from the RCF directly, which applies Census Bureau business rules to select a single address per person-year. The RCF is an internal Census Bureau product available for all states regardless of whether the state participates in LEHD for external researchers (Graham et al., 2017). The American Community Survey (ACS) identifies the study population and provides the household roster, household composition, and demographic characteristics used to reconstruct retrospective fertility measures. ACS person records are assigned PIKs via the BOC PIK Crosswalk ACS, enabling linkage to the administrative residence files. The ACS request spans 2000–2024 so early fertility-environment measures can be constructed consistently and the most recent ACS vintage is available for validation and auxiliary tabulations, but because the BOC PIK Crosswalk ACS begins in 2001, the linked ACS–administrative analytic sample begins in 2001 while the administrative residence histories extend from 2000 through 2023. Figure 1 illustrates the linkage structure across the core residence files.

Residence histories constructed from the MARF present measurement challenges. In pre-2012 years, multiple residence candidates may appear for a given individual in the same year because information is drawn from multiple administrative sources, raising ambiguity about the true residence. In 2012 and later years, a single selected address per year is recorded, which may not always

[1]For SAP request purposes, the manuscript shorthand "MARF" corresponds to BOC Master Address File Auxiliary Reference File (MAFARF) (RDG 720) 2000–2023 and BOC Master Address File–Multiple MAFIDs (RDG 18744) 2013–2021, 2023–2025. The manuscript shorthand "RCF" is requested via Longitudinal Employer-Household Dynamics (LEHD) Person-level Files Residence Components (RDG 7438), including all states.

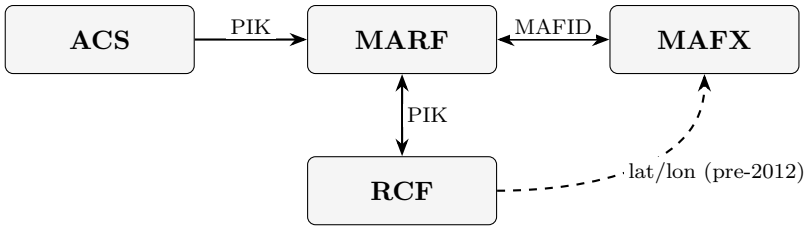


Figure 1: **Data linkage structure.** Arrows indicate links between files, labeled by the linking variable. The single-headed arrow (ACS→MARF) indicates that ACS records are linked into the residence panel via the Protected Identification Key (PIK). Double-headed arrows indicate bidirectional merge links. The dashed arrow shows the pre-2012 strategy for recovering address identifiers (MAFIDs): latitude/longitude coordinates from the RCF are used to link records to the MAFX. Together, the four files yield a longitudinal residence history for each individual in the sample.

correspond to the individual's primary residence.

Supplementary analyses use the Census Household Composition Key File (`CENSUS_HH_COMP_KEY`) for administrative parent–child links, the 2000 Decennial Census and, subject to RDC availability at the time of analysis, the 1990 Decennial Census for extended household-link observation, the SIPP through the BOC Consolidated PIK Crosswalk SIPP for validation with exact child birth timing, CPS ASEC and the CPS Fertility Supplement through their respective BOC crosswalks for additional fertility checks, and the EIF Demographic Spine and EIF Residential Information for environmental heterogeneity analyses.

The RCF is trained and validated primarily using ACS and LEHD information, so residence changes may be measured less accurately for women with weak labor-force attachment. This limitation is treated as both a source of measurement error in the main design and a motivation for the project's Criterion 5 residence-quality analyses.

The main analytic sample consists of women who meet two criteria. First, individuals must experience a qualifying residential move during the childbearing ages. Second, the move must result in a change in place-based exposure, defined as a move across relevant sub-state geographic boundaries. The analysis focuses on movers, as identification relies on comparisons among individuals who relocate to destinations with different fertility environments.

The sample is constructed using the following procedure. First, women ages 15 to 45 are identified in the ACS. We use the ACS to reconstruct fertility measures retrospectively and to characterize demographic and household circumstances observed at the survey date. Variables

drawn from the ACS include age (AGEP), sex (SEX), household composition, relationship information from the household roster (including RELSHIPP), co-resident children and their ages, and other demographic characteristics. Because the ACS does not collect a children-ever-born question after 1998, the fertility reconstruction relies on the household roster, specifically the presence, ages, and household relationships of co-resident children, rather than on a direct lifetime fertility-history question. Second, ACS records are linked to address-based administrative residence information using restricted-use person identifiers (PIKs). Through the MARF, individuals are assigned to residential addresses (MAFIDs) in each calendar year. These address records are used to construct annual residence histories that record the individual's location and associated geographic identifiers over time.

Third, residence histories are processed to identify residential moves. Moves are defined as changes in residential location that cross relevant sub-state geographic boundaries, such as counties or commuting-zone–equivalent areas, and therefore plausibly alter exposure to local fertility environments. For individuals with multiple address records in a given year, Census-recommended procedures are applied to determine residence, and alternative residence definitions are retained for robustness analyses.

Fourth, residence histories are harmonized using the RCF to resolve ambiguities and ensure consistent geography across the full sample period. For 2012 and later years, the RCF contains MAFIDs directly and is used as-is to assign annual residential addresses. For the pre-2012 period, the RCF does not contain MAFIDs but does provide latitude and longitude coordinates, which are sufficient to assign commuting-zone geography directly without requiring a MAFID. Where the requested MAFX vintages are available (2008–2026), the 2008–2011 RCF lat/lon coordinates are additionally linked to the MAFX to recover an address-level MAFID, supporting the Criterion 5 address-agreement comparisons described below. For 2000–2007, prior to the requested MAFX coverage, residence is assigned at the commuting-zone level directly from the RCF/CPR lat/lon coordinates without MAFID recovery. This strategy produces a residence history that is consistent at the commuting-zone level across the full 2000–2023 administrative residence panel, with address-level MAFID precision available from 2008 onward. Discrepancies between MARF-based and RCF-based residence assignments are documented and incorporated into robustness analyses.

Fifth, fertility histories are retrospectively reconstructed for each mover from the ACS household

roster. For each co-resident child linked to the woman through the roster, birth year is approximated as survey year minus the child's age in completed years (AGEP). Because AGEP is recorded at the interview date rather than as an exact birth date, this approximation carries approximately ± 1 year of uncertainty. Children whose inferred birth year falls before the observed move year T_i contribute to pre-move parity; children whose inferred birth year falls within a fixed window after T_i contribute to the post-move birth count. Children whose inferred birth year coincides with the move year itself are excluded from both counts to avoid misclassification around the transition date. This reconstruction yields one observation per mover rather than a repeated panel.

Finally, the analytic sample is restricted to individuals with complete linkage across data sources, non-missing key demographic variables, and a common post-move exposure window. Let S_i denote the ACS survey year and let a_i^0 denote age at move. The main estimating samples require $S_i \geq T_i + W$, where W is the chosen post-move window, so each mover is observed after the same minimum exposure length. The preferred sample restricts attention to movers whose first qualifying move occurs during the relevant fertility-risk ages, and robustness checks further restrict the sample so that the full post-move observation window falls within the core childbearing ages. Additional restrictions based on the age of the oldest co-resident child and bounded exposure horizons are used in robustness analyses to reduce undercounting of older pre-move children who may no longer be in the household at the interview date. The resulting sample consists of women of childbearing age with well-defined migration events and retrospective fertility measures suitable for implementing the mover-based empirical strategy described in the Overview.

As a supplementary measure of fertility, the analysis also draws on the Census Household Composition Key File (`CENSUS_HH_COMP_KEY`), an administrative records dataset that directly associates children up to age 18 with their co-resident adult parents via PIKs (Genadek et al., 2022). Unlike ACS-based fertility measures, which depend on survey response and household roster structure, the Household Composition Key File captures parent–child co-residence relationships in administrative records and covers a larger sample of individuals. This file is available beginning in 2016; robustness analyses comparing ACS-based and administrative-records-based fertility measures are restricted to the 2016–2023 subsample. The analysis supplements the Household Composition Key File with the 2000 Decennial Census, which provides census-recorded household relationship data for individuals age 18 or younger in 2000, extending household-link observation to earlier cohorts. Where

PIK-linked 1990 Decennial records are available at the time of analysis, this extension reaches individuals age 18 or younger in 1990.

8.3 Notation and Key Variables

The unit of analysis is mover i, observed once in ACS survey year S_i. The following variables are the key inputs to the specifications in the next section.

Residential location and moves: Let c_i^0 and c_i^1 denote the origin and destination commuting zones for mover i, constructed from the annual residence history. The first observed move year is T_i, with c_i^0 measured in $T_i - 1$ and c_i^1 measured in T_i.

Observation horizon: Let $h_i = S_i - T_i$ denote the number of calendar years between the observed move year and the ACS interview. The main estimating sample requires $h_i \geq W$, where W is the common post-move exposure window.

Retrospective fertility measures: Let $\widehat{B}_{ij}$ denote the inferred birth year of child j linked to woman i through the ACS household roster. Pre-move parity is defined as

$$P_i^{\text{pre}} = \sum_j \mathbf{1}\{\widehat{B}_{ij} < T_i\}, \tag{1}$$

and post-move childbearing within a fixed window is defined as

$$B_i^{\text{post}}(W) = \sum_j \mathbf{1}\{T_i < \widehat{B}_{ij} \leq T_i + W\}. \tag{2}$$

Children whose inferred birth year equals T_i are excluded from both counts because births in the move year cannot be cleanly assigned to the pre-move or post-move period. Because child birth years are inferred from AGEP rather than exact birth dates, these are retrospective measures with approximately ± 1 year of uncertainty around the move year.

For descriptive timing profiles, the analysis also considers coarse post-move bins such as

$$B_i^{[0,2]} = \sum_j \mathbf{1}\{T_i < \widehat{B}_{ij} \leq T_i + 2\} \tag{3}$$

and

$$B_i^{[3,5]} = \sum_j \mathbf{1}\{T_i + 2 < \widehat{B}_{ij} \leq T_i + 5\}. \tag{4}$$

Destination fertility environment: Each commuting zone c is assigned a fertility environment measure F_c, constructed from ACS data on long-term non-moving residents (excluding movers) to avoid mechanical correlation with movers' own outcomes. The change in fertility environment associated with a move is $\Delta F_i = F_{c_i^1} - F_{c_i^0}$.

Covariates: X_i^0 denotes predetermined or slow-moving characteristics measured before or at the move, including race and ethnicity, birth cohort, and flexible controls for age at move. Interview-date marital status or education, when used, are treated as robustness controls rather than baseline causal controls. The preferred specifications do not control for interview-date household composition because it may mechanically reflect the presence of children used to construct the fertility outcome. The specifications below also condition on origin fixed effects, move-year or move-cohort effects, and survey-year fixed effects.

8.4 Econometric Model

Fertility responses to migration can be represented in an annual event-time framework similar to mover designs used in other place-effects settings. Let Y_{it} denote an indicator equal to one if mover i gives birth in event year t relative to the move, where $t = 0$ denotes the move year, $t < 0$ denotes pre-move years, and $t > 0$ denotes post-move years. The causal event-time formulation is

$$Y_{it} = \alpha_t + \beta_t \Delta F_i + f_t(P_i^{\text{pre}}) + X_i^{0\prime}\delta_t + \mu_{c_i^0} + \lambda_{T_i+t} + \varepsilon_{it}, \qquad t > 0. \tag{5}$$

This formulation captures the core mover-design intuition: among women with similar origins and similar pre-move fertility histories, differences in destination fertility environments may shift annual childbearing probabilities after a move. In the linked ACS data, however, these event-time fertility histories are reconstructed retrospectively from a single interview per woman rather than observed directly in a true annual panel. Pre-move event-time outcomes can still be summarized descriptively as balance or placebo checks, but the annual causal equation is defined for post-move years so that full pre-move parity remains predetermined relative to the outcome.

In practice, the ACS implementation reconstructs approximate event-time fertility histories from a single interview per woman using the ages of co-resident children and the observed move year T_i. Rather than estimate annual β_t coefficients directly, the main analysis aggregates fertility into fixed post-move windows and estimates a retrospective mover comparison. The main specification relates post-move childbearing within a common exposure window to the change in fertility environment between origin and destination:

$$B_i^{\text{post}}(W) = \alpha + \beta \Delta F_i + f(P_i^{\text{pre}}) + X_i^{0\prime}\delta + \mu_{c_i^0} + \lambda_{T_i} + \nu_{S_i} + \varepsilon_i. \tag{6}$$

The coefficient β captures how post-move childbearing within window W varies with the fertility environment of the destination, conditional on pre-move parity, predetermined characteristics, origin location, move year, and survey year. Conditioning flexibly on P_i^{pre} is essential because the reconstructed history allows women with similar prior childbearing histories to be compared even though fertility is observed retrospectively from a single ACS interview.

Descriptive timing profiles: To provide a coarse analogue to the conceptual annual event-time profile, the analysis also estimates versions of the main retrospective mover specification using $B_i^{[0,2]}$ and $B_i^{[3,5]}$ as outcomes. These coefficients are presented as descriptive timing profiles rather than year-specific event-study effects, because the ACS reconstruction does not support exact annual assignment of births relative to the move year.

Identification and interpretation: Identification in both the conceptual event-time representation and the implemented retrospective specification relies on comparisons among women from similar origins and with similar pre-move parity who relocate to destinations with different fertility environments. The key identifying assumption is conditional exogeneity of ΔF_i: after conditioning on P_i^{pre}, predetermined characteristics X_i^0, origin fixed effects, move-year or move-cohort effects, and survey-year fixed effects, destination fertility environments are as good as random with respect to unobserved determinants of post-move childbearing. Because origin fixed effects absorb the origin component $F_{c_i^0}$, identification comes from variation in destination fertility environments $F_{c_i^1}$ among movers from the same origin. The leading threat is selective destination choice based on anticipated fertility plans. For example, a woman who intends to have additional children and

deliberately chooses a high-F destination for reasons not captured in X_i^0 would violate the identifying assumption. The design is therefore best understood as a single-index mover specification: rather than estimate unrestricted destination place effects, it asks whether moves to destinations with higher measured fertility environments are associated with higher post-move fertility. The design therefore emphasizes rich predetermined controls, origin fixed effects, and robustness checks across alternative residence definitions and sample restrictions (Finkelstein et al., 2016, 2021).

As a direct probe of this threat, the analysis also examines a robustness subsample of movers who relocate into commuting zones during years experiencing large, discrete positive employment shocks. These shocks are constructed from LBD-identified establishment births or expansions at the commuting-zone-by-industry-by-year level and then aggregated to a destination commuting-zone-by-year indicator for the primary ACS subsample. A more restrictive industry-matched version retains the industry dimension and uses observed industry of employment where available. In SIPP validation analyses, this classification can be supplemented with self-reported job-related reasons for moving. For movers entering these shock CZ-years, the relocation decision is more likely to be tied to employment opportunities than to unobserved fertility preferences, so a similar estimated response to ΔF_i in this subsample would reduce the concern that the full-sample estimates are driven entirely by fertility-motivated destination sorting.

This exercise is not an instrumental-variables design and does not impose an exclusion restriction. Employment shocks may also affect household income, spousal employment, housing costs, and local labor market conditions, all of which may affect fertility directly. The subsample analysis should therefore be interpreted as a robustness exercise that plausibly reduces sorting on fertility preferences at the cost of introducing a bundled local economic shock that cannot be fully separated from the fertility-environment channel. This analysis is described further below.

Births assigned to the move year itself are excluded from the main estimand because annual residence histories and AGEP-based birth-year inference cannot determine whether they occurred before or after the move. The resulting estimates should therefore be interpreted as the relationship between destination fertility environments and childbearing realized after the transition year, not as an exact event-time response in the move year itself ($t = 0$).

Inference: All estimates are produced within the restricted-use Census Bureau environment and comply with disclosure avoidance requirements. Standard errors are discussed in the implementation details below.

8.5 Implementation Details

This section describes key practical choices involved in implementing the econometric specifications described above.

Construction of fertility environments: Local fertility environments are constructed using ACS data on fertility behavior among long-term residents. For each geographic unit c, the fertility environment measure F_c is computed using women who do not experience a residential move during a specified pre-period. Fertility environments are based on age-specific fertility rates and related household composition measures derived from the ACS. To avoid mechanical correlation with movers' outcomes, fertility environment measures are constructed excluding movers and using lagged or pre-period data.

Definition of residential moves: Residential moves are defined as changes in sub-state geographic location across consecutive years. Moves that do not cross meaningful geographic boundaries (e.g., within-commuting-zone moves) are excluded from the main analysis, as they do not generate a change in the local fertility environment and are uninformative for the main estimand.

Multiple residential moves: The main analysis restricts attention to each woman's first observed residential move, which defines the move year T_i, origin c_i^0, and destination c_i^1. Restricting to the first move avoids complications arising from multiple exposure changes and yields a common interpretation of the retrospective post-move window. Women who move more than once during the observation window contribute to the main analysis through their first move only. Robustness checks compare results to a subsample restricted to single-move women and to alternative residence-history constructions.

Alternative fertility measures: The primary fertility measures are reconstructed from the ACS household roster using co-resident children linked to the woman through the roster and the

child's age in completed years. As a robustness check, the analysis also constructs fertility measures using the Census Household Composition Key File (`CENSUS_HH_COMP_KEY`), which directly associates children up to age 18 with their co-resident adult parents in administrative records via PIKs. Unlike ACS-based measures, these do not rely on survey response and are available for a larger set of individuals. Because the Household Composition Key File is available beginning in 2016, this robustness comparison is restricted to the 2016–2023 subsample. The analysis supplements the Household Composition Key File with the 2000 Decennial Census, which provides census-recorded household relationship data for individuals age 18 or younger in 2000, extending household-link observation to earlier cohorts. Where PIK-linked 1990 Decennial records are available at the time of analysis, this extension reaches individuals age 18 or younger in 1990. Consistency between ACS-based and administrative-records-based results would strengthen confidence that the main estimates are not driven by measurement limitations of the ACS.

SIPP and CPS validation: Where available in the relevant SIPP panels and modules, SIPP provides exact birth month and year for biological children, allowing the same retrospective mover design to be implemented without AGEP-based timing uncertainty. The proposal uses SIPP as a validation and measurement-error calibration dataset rather than as the primary analysis sample: parity-within-window estimates, coarse dynamic profiles, and pre-move parity classifications from the ACS reconstruction will be compared against their SIPP analogues. Supplementary analyses also use CPS ASEC household composition as an annual check on changes in co-resident children and, for the subset linkable through CPS rotation overlap (BOC Consolidated PIK Crosswalk CPS, RDG 2593, available 1973–2025), the CPS Fertility Supplement to compare the ACS reconstruction with direct fertility questions. PIK-linked CPS Fertility Supplement validation therefore covers the full requested 1998–2022 biennial waves.

Employer-side shock subsample: As described in the identification discussion above, the analysis constructs a robustness subsample of movers whose relocation is more likely to be tied to a local employment shock than to fertility plans. First, the Longitudinal Business Database (LBD) is used to identify large, discrete positive employment events, such as new establishment births or existing-establishment expansions, at the commuting-zone-by-industry-by-year level. This ap-

proach is motivated by the large-plant-opening literature, including Greenstone et al. (2010), but does not replicate a winner-runner-up plant siting design. The LBD information is used to classify destination commuting-zone-years and commuting-zone-by-industry-by-year cells; individual movers are not linked to specific establishments. Second, movers in the ACS analytic sample are flagged as shock-exposed if they relocate into a flagged commuting-zone-year within a short event window around the shock year. The primary version defines shock exposure using destination commuting zone and move year only. A more restrictive industry-matched version, implemented among employed movers with observed survey-year industry, additionally requires that the mover's reported industry of employment match the shock industry. In SIPP waves with reason-for-move information, this classification is supplemented with self-reported job-related reasons for moving. Third, the main retrospective specification is re-estimated within the shock-exposed subsample as a robustness check on the full-sample estimate of β.

This subsample should be interpreted with two caveats. First, large employment shocks in male-dominated industries, such as some manufacturing plant openings, may disproportionately draw male workers, so shock-exposed women movers may be overrepresented among women who relocate because a spouse or partner took a position tied to the shock rather than because of their own employment. This may reduce direct sorting on the woman's own fertility preferences, but it does not eliminate household-level selection into migration and it means the subsample is not representative of movers generally. Any resulting estimate therefore describes this tied-mover population specifically. Second, the number of movers who can be classified as entering a qualifying employment-shock commuting-zone-year within the ACS and SIPP samples is expected to be modest, so results from this subsample are best treated as an illustrative, corroborating check on the main estimates rather than as a precisely estimated stand-alone result.

Exposure window and sample restrictions: The main analysis is conducted using a fixed post-move exposure window W. The baseline specification will compare conservative and broader windows, such as $W = 3$ and $W = 5$, and requires $h_i \geq W$ so that all movers are observed after the same minimum exposure length. Births assigned to the move year itself are excluded from the main sample. Because the ACS outcome is reconstructed from co-resident children observed at the interview date, robustness analyses also restrict the elapsed horizon to bounded ranges and, in

alternative specifications, include flexible controls for h_i to assess whether results are sensitive to differences in the time between the move and the ACS interview. Additional restrictions based on age at move, the age of the oldest co-resident child, and related household composition checks are used in robustness analyses to reduce the likelihood that older pre-move children have already left the household by the ACS interview.

Exposure-duration dose response: As an additional robustness and heterogeneity exercise, the analysis tests whether the estimated relationship between destination fertility environments and post-move childbearing varies with the elapsed post-move horizon. Specifically, the specification is augmented with an interaction between ΔF_i and $h_i = S_i - T_i$, the number of calendar years between the move and the ACS interview. This exercise is motivated by exposure-duration designs in the neighborhood-effects literature (Chetty et al., 2016; Chetty and Hendren, 2018a), but it is interpreted cautiously because the ACS fertility outcome is reconstructed retrospectively from one interview rather than observed as an annual panel outcome. The interaction is therefore treated as a dose-response and robustness check, not as a stand-alone causal timing design.

Controls and fixed effects: All preferred specifications include origin commuting-zone fixed effects, move-year or move-cohort effects, and survey-year fixed effects. Baseline covariates are predetermined or slow-moving characteristics, including race and ethnicity, birth cohort, and flexible age-at-move controls. Pre-move parity enters separately through $f(P_i^{\text{pre}})$. Interview-date marital status or education may be added as robustness controls where useful, but interview-date household composition is excluded from the preferred causal specification because it may mechanically reflect the fertility outcome itself. Individual fixed effects are not used because each mover contributes one retrospective observation rather than repeated annual fertility outcomes.

Environmental mechanism heterogeneity: The EIF Residential Information is used in supplementary analyses to measure destination environmental quality, including pollution exposure and temperature-related conditions, once the destination residence history has been assigned. These analyses replace or augment ΔF_i with changes in environmental exposure to assess whether fertility responses to migration differ across cleaner or more hazardous destination environments.

Weights and inference: The primary specifications are estimated without ACS survey weights, consistent with the mover-based literature, because identification relies on comparisons among movers with similar origins and observed characteristics rather than on population representativeness. Robustness checks using ACS weights are conducted where feasible. Standard errors are clustered at the origin commuting-zone level in the baseline specification to allow for correlated shocks among movers leaving the same place. Additional inference checks will consider alternative clustering structures, including two-way clustering by origin and destination commuting zones.

8.6 Feasibility

The project is feasible within the RDC environment using existing Census Bureau data products and established linkage procedures. The ACS provides a large annual probability sample with rich demographic and household information, while Census administrative residence files allow construction of longitudinal residence histories. The use of PIKs and existing crosswalks enables reliable linkage across the residence panel, ACS, and supplementary validation files.

The proposed analysis relies on standard data structures and methods that have been implemented in prior RDC-approved projects, including the construction of residence histories, identification of migration events, and estimation of mover-based comparison models. Preliminary tabulations indicate that the number of movers between distinct fertility environments is sufficiently large to support the proposed analyses, even after imposing age, linkage, and common-window restrictions.

Computational demands are moderate and compatible with RDC computing resources. The project does not require the creation of new linkages beyond those already supported within the RDC infrastructure. The LBD robustness component requires only the construction of aggregate commuting-zone-by-industry-by-year shock measures and a merge of those aggregate measures to ACS and SIPP movers by destination commuting zone, move year, and, in the stricter version, industry. It does not require linking individual movers to specific firms or establishments.

All outputs will be submitted for Census Bureau disclosure review prior to release. The primary outputs of this project are regression coefficients and standard errors from retrospective mover comparison models and coarse post-move timing specifications estimated on the national sample of movers. These types of outputs typically pass disclosure review under standard Census Bureau

rules, as they do not reveal information about specific individuals. The project does not request sub-national tabulations of individual-level microdata, which are the most common source of disclosure failures. The project does not intend to request geographic-area estimates below the state level, and in most cases will not request sub-state estimates at all. Commuting-zone and origin fixed effects are used internally in the estimation models to absorb geographic heterogeneity, but the fixed-effect coefficients themselves are not requested as outputs; only aggregate coefficients such as the coefficient on ΔF_i and related summary statistics are requested for release. This design avoids concerns related to Geographic Areas with Small Populations (GASP). For the employer-side shock robustness component, industry categories from the LBD will be aggregated as needed to meet Census Bureau disclosure concentration-ratio and minimum sample-count requirements before any industry-level shock definitions or related estimates are requested for release. The project team will consult with RDC staff during the analysis phase to confirm appropriate aggregation levels and to flag any outputs that may require special review, and will follow Census Bureau disclosure avoidance guidelines throughout. A full description of anticipated modeling and tabular output, keyed to the SAP portal's Requested Output field, is provided as a separate attachment.

8.7 Expected Outcomes

The project will produce three main types of outcomes.

First, it will generate estimates of how fertility and family size outcomes respond to changes in place-based fertility environments following residential moves. These estimates will shed light on the role of local social environments and norms in shaping fertility behavior over the life course.

Second, the project will provide evidence on the timing of fertility responses to migration within fixed post-move windows, distinguishing near-term adjustments from later childbearing. This contributes to the literature on place-based effects by extending mover-based designs to fertility outcomes.

Third, the project will produce internal documentation evaluating the construction and use of longitudinal residence histories, household composition measures, and administrative residence assignments in Census Bureau data. This includes comparisons of alternative residence definitions and assessments of how measurement choices affect estimates of family-related outcomes. These materials will be shared within the Census Bureau as technical memoranda and will inform future

research using Census administrative and survey data.

Overall, the project advances understanding of how residential mobility interacts with family formation decisions while simultaneously contributing to improved use and interpretation of Census Bureau data products.

9 Benefits

The predominant purpose of this project is to increase the utility of Title 13, Chapter 5 data by generating new population estimates and improving understanding of administrative residence history measurement in Census Bureau data products. The project meets two Census Bureau benefit criteria as described below.

Criterion 5: Understanding and/or improving the quality of data produced through a Title 13, Chapter 5 survey, census or estimate

The project will improve the Census Bureau's understanding of the quality of residence histories constructed from administrative records for individuals observed in the American Community Survey (ACS). Fertility outcomes and household composition as recorded in the ACS reflect conditions at the time of survey response and are accurately measured for that survey round. However, constructing the longitudinal residence histories required for this project's identification strategy involves linking ACS respondents to residence-history sources—the Master Address Administrative Reference file (MARF), the Residence Candidates File (RCF), and the EIF Residential Information—whose address assignments may differ from the physical address recorded in the ACS itself. The ACS contains a MAFID corresponding to the physical address at which the survey was completed; the MARF and RCF, drawn from administrative records, may assign different addresses than the ACS-recorded address for the same individual. Sullivan and Genadek (2024) document baseline ACS–MARF agreement rates across the full ACS population, but whether MARF, RCF, and EIF residence assignments are systematically more or less consistent with ACS-recorded addresses for women with co-resident children has not been examined.

The project will assess residence history data quality along two dimensions. First, it will extend the framework of Sullivan and Genadek (2024), who document that approximately 75–80

percent of ACS respondents have a MAFID recorded in the ACS that agrees with their MARF-assigned address. This project asks whether match rates differ systematically for women with and without co-resident children—testing whether the presence of children is associated with more or less consistent administrative records of residence. The comparison covers the MARF, the RCF, and the EIF Residential Information, which differ in important ways. Prior to 2012, the MARF could assign multiple MAFIDs to a single individual; beginning in 2012, when multiple address candidates exist, one is selected at random, introducing address uncertainty into post-2012 records. The RCF uses a statistical model trained on ACS data to assign a single best-guess residence for each individual; because the model was designed to perform well for employed persons, its predictions may be less accurate for women of childbearing age who are out of the labor force (Graham et al., 2017). The EIF Residential Information provides a third residence-history source built from administrative and commercial data, allowing the analysis to assess whether conclusions about residence assignment quality are consistent across MARF, RCF, and EIF. Documenting how fertility and household composition moderate consistency with ACS-recorded addresses for each source separately will clarify for which subpopulations and time periods each residence-history source is most reliable.

Where the Census Household Composition Key File permits identification of co-resident parent–child pairs (2016–2023 subsample), the analysis will further classify pairs by the consistency of their administrative residence records with ACS-recorded addresses: (1) both parent and child records agree with the ACS address; (2) the parent record agrees but the child record does not; (3) the child record agrees but the parent record does not; and (4) neither record agrees. This classification will characterize the conditions under which administrative sources consistently place co-resident parents and children at the same address, and identify where residence histories are most likely to contain errors for fertility-relevant household configurations. Where data permit, the analysis will be extended to fathers. The richness of this parent–child matching analysis presents an opportunity for a standalone publication extending Sullivan and Genadek (2024). Where 2000 Decennial records are linked via the BOC PIK Crosswalk Census Unedited File (CUF), and, subject to availability of PIK-linked 1990 Decennial records, to earlier cohorts, the household-relationship matching analysis provides a longer window to characterize how fertility and household configuration relate to residence assignment quality across Census and administrative data vintages.

Second, the project will specifically examine cases in which fertility and household structure create ambiguity in residence assignment. A particularly important example involves children of separated or divorced parents, who may be associated with two different residential addresses simultaneously. In pre-2012 MARF records, multiple MAFIDs could be assigned to the same individual, potentially revealing this split. In post-2012 records, however, a single MAFID is selected at random from the available candidates, which may make such dual-address situations harder to detect—and potentially introduce greater error for households with children across two residences. In such cases, the child's presence may be recorded in one parent's household in one data source and in the other parent's household in another source, creating inconsistencies in household composition measures that directly affect fertility indicators. The project will document the prevalence of such cases across both periods and their implications for residence history quality. These findings will be prepared as a technical memorandum to remain on the Census Bureau server and will inform future Census Bureau research and data product development.

Criterion 11: Preparing estimates of population and characteristics of population as authorized under Title 13, Chapter 5

The project will produce new estimates of the relationship between residential mobility and fertility and family size decisions for women of childbearing age in the United States. Existing Census Bureau publications and data products do not contain estimates of how a change in place of residence causally affects subsequent fertility behavior, conditional on prior fertility and demographic characteristics. The project will generate such estimates using restricted-use ACS microdata requested for all 50 states and the District of Columbia from 2000 to 2024.

Specifically, the project will estimate retrospective mover comparison models relating post-move childbearing within fixed exposure windows to the fertility environment of the destination commuting zone, among women ages 15 to 45 who experience a residential move. Fertility outcomes will be reconstructed from the ACS household roster to measure pre-move parity, post-move births, and parity progression. Estimates will be produced for the full sample of movers and for subgroups defined by age at move, prior parity, and characteristics of the move. These coefficient estimates characterize the fertility behavior of a large and demographically important subpopulation—women of childbearing age who relocate within the United States—in a way that no existing Census Bureau

product does.

The estimates will be submitted for publication in a peer-reviewed economics journal. The project anticipates producing one research paper, targeting a top general-interest economics journal such as the *American Economic Review*, *Journal of Political Economy*, or *Quarterly Journal of Economics*. Prior to journal submission, the paper will be submitted to the Census Bureau's Center for Economic Studies (CES) Working Paper series. Results will also be presented at major economics and health economics conferences, including the AEA/ASSA Annual Meeting, the Western Economic Association International (WEAI) Annual Conference, the Southern Economic Association (SEA) Annual Meeting, the Association for Public Policy Analysis and Management (APPAM) Fall Research Conference, and the American Society of Health Economists (ASHEcon) Conference. Publication in a top venue will make the estimates widely available to researchers, policymakers, and the public, extending the reach and utility of Census Bureau data.

Deliverables

Criterion 5: The project will produce one or more internal technical memoranda documenting findings on the quality of administrative residence histories relative to ACS-recorded addresses, including: fertility-conditioned agreement rates between ACS-recorded addresses and MARF-, RCF-, and EIF-assigned addresses; parent–child matching categories derived from the Census Household Composition Key File and, where 2000 Decennial records are linked and subject to availability of PIK-linked 1990 Decennial records, extended to earlier cohorts; and the prevalence of ambiguous residence assignment cases (such as children of separated parents appearing at multiple addresses). All memoranda will remain on the Census Bureau server.

Criterion 11: The project will produce one peer-reviewed research paper presenting estimates of the causal effect of residential moves on fertility and family size decisions for women of childbearing age. The paper will be submitted to the Census Bureau's CES Working Paper series and to a top general-interest economics journal following Census Bureau disclosure review. Results will be presented at major economics and health economics conferences (AEA/ASSA, WEAI, SEA, APPAM, ASHEcon) prior to publication.

References

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2, Part 2):S279–S288.

Chetty, R. and Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *Quarterly Journal of Economics*, 133(3):1163–1228.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Ciliberto, F., Miller, A. R., Nielsen, H. S., and Simonsen, M. (2016). Playing the fertility game at work: An equilibrium model of peer effects. *International Economic Review*, 57(3):827–856.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2023). The economics of fertility: A new era. In *Handbook of the Economics of the Family*, volume 1, pages 151–254. Elsevier.

Ferrara, E. L., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of geographic variation in health care: Evidence from patient migration. *The quarterly journal of economics*, 131(4):1681–1726.

Finkelstein, A., Gentzkow, M., and Williams, H. (2021). Place-based drivers of mortality: Evidence from migration. *American Economic Review*, 111(8):2697–2735.

Genadek, K., Stevenson, A., and Sanders, J. (2022). Measuring US fertility using administrative data from the Census Bureau. *Demographic Research*, 47(2):37–58.

Graham, M. R., Kutzbach, M. J., and Sandler, D. H. (2017). Developing a residence candidate file for use with employer-employee matched data. CES Working Paper CES-WP-17-40, U.S. Census Bureau, Center for Economic Studies. Accessed: 2026-03-07.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598.

Kearney, M. S. and Levine, P. B. (2015). Media influences on social outcomes: The impact of mtv's *16 and Pregnant* on teen childbearing. *American Economic Review*, 105(12):3597–3632.

Kearney, M. S., Levine, P. B., and Pardue, L. (2022). The puzzle of falling us birth rates since the great recession. *Journal of Economic Perspectives*, 36(1):151–176.

Lalive, R. and Zweimüller, J. (2009). How does parental leave affect fertility and return to work? evidence from two natural experiments. *Quarterly Journal of Economics*, 124(3):1363–1402.

March of Dimes (2026a). Fertility rate: District of columbia, 2014–2024. PeriStats. Reports a fertility rate of 39.8 births per 1,000 women ages 15–44 in the District of Columbia in 2024. Accessed May 3, 2026.

March of Dimes (2026b). Fertility rate: South dakota, 2014–2024. PeriStats. Reports a fertility rate of 66.7 births per 1,000 women ages 15–44 in South Dakota in 2024. Accessed May 3, 2026.

Milligan, K. (2005). Subsidizing the stork: New evidence on tax incentives and fertility. *Review of Economics and Statistics*, 87(3):539–555.

Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the united states. *Journal of Economic Perspectives*, 25(3):173–96.

Schaller, J. (2016). Booms, busts, and fertility: Testing the becker model using gender-specific labor demand. *The Journal of Human Resources*, 51(1):1–29.

Sullivan, J. and Genadek, K. (2024). Using the Census Bureau's Master Address File for migration research. ADEP Working Paper 2024-01, U.S. Census Bureau, Associate Directorate for Economic Programs.