A Cohort-by-Age Approach to Fertility Policy Evaluation

Mi Luo* Anson Zhou†

March 17, 2026
[Click here for the latest version]

Abstract

Fertility policies may shift birth timing, increase completed fertility, or both. Distinguishing these effects typically requires observing treated cohorts through the end of childbearing, often decades after implementation. We develop a framework decomposing causally identified cohort-by-age fertility responses into quantum (completed fertility) and tempo (timing) components, enabling extrapolation under testable restrictions. Applying the method to Alaska's Permanent Fund Dividend using CPS data from 1976–2025, we find substantial tempo effects: the policy accelerated childbearing toward younger ages, inflating period measures. Early treated cohorts show completed fertility increases of approximately 0.1 births per woman, but effects attenuate for more recent cohorts.

JEL classification: J13, J18, H31, C53
Keywords: Fertility, policy evaluation, completed fertility

*Guanghua School of Management, Peking University
†Faculty of Business and Economics, The University of Hong Kong.

1 Introduction

Evaluating fertility policies presents a fundamental timing problem. The outcome of interest in many cases—completed fertility—is observed only after treated cohorts finish childbearing, typically decades after a policy begins. Researchers instead rely on period measures such as the total fertility rate (TFR), which are available immediately but conflate two distinct responses: changes in the number of children women ultimately have (quantum) and shifts in when they have them (tempo). A policy that accelerates childbearing without changing family size raises period TFR in the short run and lowers it later. Conversely, a policy that increases completed fertility but induces delay may show muted or even negative period effects initially. Without separating these channels, standard approaches risk mistaking temporary timing shifts for permanent fertility changes.

This distinction has taken on new urgency. Fertility has fallen dramatically worldwide, from about five births per woman in 1950 to just above two today, with two-thirds of the world's population now living in countries with below-replacement fertility (Geruso and Spears, 2026). Governments have responded with a range of pro-natalist interventions—cash transfers, parental leave expansions, childcare subsidies—yet lack reliable methods to assess whether these policies increase completed fertility or merely shift its timing. The stakes are high: a policy that appears successful based on short-run TFR gains may leave long-term demographic trajectories unchanged if the observed response reflects only accelerated childbearing. As fertility decline increasingly shapes fiscal planning, labor force projections, and social insurance design, policymakers need tools that distinguish genuine increases in family size from temporary timing shifts.

This paper develops a framework for decomposing causally identified fertility responses into quantum and tempo components, enabling researchers to assess completed fertility effects before treated cohorts finish childbearing. The method applies to any setting where cohort-by-age fertility responses can be estimated—whether from difference-in-differences, regression discontinuity, or synthetic control designs evaluating cash transfers, parental leave, childcare subsidies, or other interventions. The key insight is that if treatment effects decompose additively into cohort effects (quantum) and age effects (tempo), the structure can be estimated from observed data and used to project completed fertility for cohorts still in their childbearing years. The additive restriction is testable, providing a diagnostic for when such extrapolation is warranted.

The framework proceeds in four steps, developed formally in Section 3. First, we esti-

mate fertility responses for each cohort-age cell using standard causal inference methods. Second, we decompose these cell-level estimates into additive cohort and age fixed effects, where cohort effects capture quantum and age effects capture tempo. Third, we extrapolate cohort effects for cohorts not yet fully observed using local polynomial regression. Fourth, we aggregate these components to construct completed fertility projections and period TFR paths. The assumptions underlying extrapolation—additivity and smooth evolution of cohort effects—are explicit and can be evaluated against observed data.

We demonstrate the framework using the Alaska Permanent Fund Dividend (APFD), an annual unconditional cash transfer to Alaska residents that began in 1982. Using micro-level data from the Current Population Survey from 1976 to 2025, we estimate cohort-by-age fertility responses via difference-in-differences with the rest of the United States as the control group, then apply the decomposition. Section 5 presents the complete empirical analysis.

The results illustrate why the quantum-tempo distinction matters for policy evaluation. The APFD generated substantial tempo effects, accelerating childbearing toward younger ages. Women in their early twenties show fertility increases of approximately 1.7 percentage points relative to the baseline, offset by reductions at older ages. These timing shifts inflated period TFR substantially during the mid-1980s. Quantum effects are more modest and concentrated among specific cohorts: women born in the late 1960s, who experienced the policy throughout most of their childbearing years, show the largest completed fertility gains. Earlier cohorts, exposed to the policy only during their remaining fertile years, show smaller effects. For cohorts entering childbearing today, extrapolated effects on completed fertility are near zero, suggesting that the APFD's long-run demographic impact has largely materialized.

We conduct robustness checks along two dimensions. First, we evaluate the identifying assumptions underlying our decomposition. A test for cohort-by-age interaction yields no evidence that timing shifts vary systematically across cohorts (p-value = 0.67), and the two-way decomposition explains nearly half the variation in cell-level treatment effects (R^2 = 0.47). Appendix B implements an alternative decomposition following Heckman et al. (1998) that imposes endpoint restrictions on age effects rather than relying solely on additivity; the estimated period effects are close to zero, supporting the baseline specification.

Second, we address the concern that parallel trends may fail if Alaska differs systematically from other states in ways that affect fertility trajectories. Appendix A implements a synthetic control approach that constructs a data-driven counterfactual by weighting donor

states to match Alaska's pre-treatment fertility rates and demographic composition. The synthetic control estimates corroborate the main findings: the correlation between difference-in-differences and synthetic control cell-level estimates is 0.89, and both methods recover the same tempo pattern of accelerated childbearing among younger women.

Related Literature

This paper contributes primarily to the methodology of fertility policy evaluation. To our knowledge, it is the first to decompose causal treatment effects estimated at the cohort-age level into quantum and tempo components, and to use these structural estimates to project policy effects on completed fertility.

A large empirical literature estimates effects of pro-natalist policies on period fertility measures, including cash transfers (Milligan, 2005; Cohen et al., 2013; González, 2013; Raute, 2019), parental leave (Lalive and Zweimüller, 2009; Dahl et al., 2016), and childcare provision (Bauernschuster and Schlotter, 2015; Givord and Marbot, 2015). These studies provide credible causal estimates of short-run responses but generally cannot distinguish timing shifts from completed fertility changes. The few studies that observe completed fertility do so by examining historical policies with sufficient follow-up (Mumford, 2017; Hart and Rønsen, 2019), sacrificing timeliness and often external validity. The framework developed here bridges this gap, extracting quantum and tempo components from the same data used for period analysis.

The decomposition builds on the demographic literature distinguishing period from cohort fertility. Bongaarts and Feeney (1998) proposed adjusting period TFR for tempo distortions using observed changes in mean age at childbearing, an approach influential in formal demography (Sobotka, 2004; Goldstein et al., 2009). The Bongaarts-Feeney adjustment corrects aggregate trends under assumptions about the uniformity of timing shifts across ages and parities. The present framework shares the goal of separating timing from quantum but differs in two important respects. First, it is designed for causal inference rather than descriptive demography, estimating cohort and age effects from experimentally or quasi-experimentally identified treatment effects rather than from observed fertility trends. Second, the restrictions it imposes are testable within the policy evaluation context, as demonstrated in Section 5.7. The marriage of causal identification with structural decomposition is, to our knowledge, novel.

Methodologically, the additive decomposition of fertility responses into cohort and age

components parallels the influential framework of Abowd et al. (1999) for decomposing wages into worker and firm fixed effects. Both approaches impose additive separability to parse variation in an outcome into distinct components with substantive interpretations—in AKM, portable worker skills versus firm-specific wage premia; here, permanent fertility responses (quantum) versus timing patterns (tempo). The identification strategies share a common structure: AKM identifies worker and firm effects through workers who move between employers, while the present framework identifies cohort and age effects through the Lexis diagram structure in which different cohorts are observed at the same age in different calendar years. A key lesson from the AKM literature is that limited mobility can bias variance estimates and correlations between components (Andrews et al., 2008; Bonhomme et al., 2019; Kline et al., 2020). The analogous concern here is that cohorts observed over few ages yield noisy cohort effect estimates, which motivates the precision-weighting approach in Section 3. Unlike AKM applications, which typically characterize the joint distribution of existing worker-firm matches, this paper's primary goal is extrapolation—projecting cohort effects forward to assess completed fertility before it is observed. This extrapolation objective, combined with the policy evaluation context, distinguishes the present contribution from the wage decomposition literature.

More broadly, the paper relates to the challenge of inferring long-run outcomes from short-run data. Structural life-cycle models address this by specifying preferences, constraints, and dynamic optimization, then simulating forward (Keane and Wolpin, 1997; Francesconi, 2002; Gayle and Miller, 2012). This approach provides rich interpretation but requires functional form assumptions that are difficult to validate. At the other extreme, reduced-form studies report period effects without extrapolation, avoiding assumptions at the cost of leaving the long-run question unanswered. The framework here occupies intermediate ground. It imposes enough structure to separate quantum from tempo and project completed fertility, while keeping assumptions transparent, parsimonious, and empirically testable. Section 4 provides a detailed comparison of these alternative approaches.

The application to the APFD contributes to a growing literature on the behavioral effects of this policy. Prior work has examined effects on consumption (Hsieh, 2003), labor supply (Jones and Marinescu, 2022), crime (Watson et al., 2020), and health at birth (Chung et al., 2016). Most relevant to the present study, Yonzan et al. (2024) use synthetic control methods to estimate that the APFD increased fertility rates by approximately 13 percent in the years immediately following its introduction, with effects concentrated among women over age 20

and no significant change in adolescent fertility or abortion rates. Their findings suggest the increase reflected planned births rather than unintended pregnancies. However, their analysis focuses on short-run period effects (1982–1988) and does not distinguish between tempo shifts and completed fertility changes. The present paper builds on their estimates by decomposing observed period effects into timing and quantum components, showing that a substantial portion of the fertility increase reflects accelerated childbearing rather than permanent changes in family size. This distinction has direct policy relevance: tempo effects imply that period-based evaluations may overstate the long-run demographic impact of cash transfer programs.

Outline

The remainder of the paper is organized as follows. Section 2 establishes notation and formalizes the extrapolation problem that motivates the decomposition. Section 3 develops the framework in four steps: estimating cohort-age responses, decomposing into quantum and tempo, extrapolating cohort effects, and aggregating to policy-relevant measures. Section 4 situates the approach within the landscape of alternatives, discussing tradeoffs relative to waiting for completed fertility, relying on period TFR, tempo-adjusted TFR, and structural models. Section 5 applies the framework to the APFD, presenting estimates of tempo effects, quantum effects, extrapolated completed fertility, and specification tests that validate the additive structure. Section 6 concludes. Appendix A presents the synthetic control robustness analysis, and Appendix B discusses alternative approaches for relaxing our identifying assumptions.

2 Setup

This section establishes notation, defines the primitive objects of interest, and formalizes the extrapolation problem that motivates the decomposition framework developed in Section 3.

2.1 Notation and Definitions

Let c denote birth cohort (year of birth), $j \in \{0, \ldots, J\}$ denote age, and t denote calendar year. These three dimensions are linked by the accounting identity $t = c + j$. Fertility occurs over ages $j \in \{J_{\text{start}}, \ldots, J_{\text{end}}\}$, corresponding to the childbearing years. We follow the conventional definition and set $J_{\text{start}} = 15$ and $J_{\text{end}} = 44$.

Define *cohort-by-age fertility* as

$$f_{j,c}, \tag{1}$$

the probability that a woman born in cohort c gives birth at age j, equivalently in calendar year $t = c+j$. This object forms the building block from which all standard fertility measures are constructed.

Consider a policy implemented in year t_0. Let $f^0_{j,c}$ denote the counterfactual fertility schedule that would have prevailed absent the policy. The *cohort-by-age fertility response* is

$$\Delta f_{j,c} = f_{j,c} - f^0_{j,c}, \tag{2}$$

measuring the causal effect of the policy on fertility for women of cohort c at age j.

We take the matrix $\{\Delta f_{j,c}\}$ as the primitive object for policy evaluation. Standard fertility measures are aggregates of these responses. The *Completed Fertility Rate* (CFR) for cohort c sums fertility across all childbearing ages:

$$\text{CFR}_c = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} f_{j,c}. \tag{3}$$

The CFR measures the total number of children born to the average woman in cohort c over her lifetime.

The *Total Fertility Rate* (TFR) in calendar year t sums age-specific fertility rates across the cohorts present in that year:

$$\text{TFR}_t = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} f_{j,\,t-j}, \tag{4}$$

where women of age j in year t belong to cohort $c = t - j$. The TFR is a synthetic measure representing the number of children a hypothetical woman would have if she experienced the age-specific fertility rates observed in year t throughout her childbearing years.

The *Crude Birth Rate* (CBR) in year t is a population-weighted average of $f_{j,\,t-j}$ across ages, incorporating the age structure of the female population. Any policy effect on CFR, TFR, CBR, or other aggregate measures is fully determined by the matrix of cohort-age responses $\{\Delta f_{j,c}\}$.

2.2 The Value of Cohort-by-Age Responses

CFR, TFR, and CBR answer fundamentally different questions, and conflating them leads to misleading policy conclusions. This point is central to the motivation for this paper.

The TFR is the most commonly reported fertility measure, prized for its timeliness and interpretability. It conflates timing shifts with changes in completed fertility, however. A policy that induces women to have children earlier without changing their lifetime number of births will raise the TFR in the short run, as births are pulled forward, and lower it later, as the women who accelerated childbearing have fewer remaining births. Evaluated at any single point in time, such a policy appears to affect fertility even though completed fertility is unchanged. Conversely, a policy that increases completed fertility but induces delay will temporarily depress the TFR. Evaluating policy based on TFR alone therefore risks mistaking tempo effects for quantum effects, precisely the concern raised by Geruso and Spears (2026).

The CFR avoids this conflation by tracking actual completed fertility for each cohort. It answers the question most relevant for long-run demographic projections and population policy: how many children did women in a given cohort ultimately have? The CFR is observed only after a cohort finishes childbearing, however, typically at age 44 or later, making it unsuitable for timely policy evaluation. A policy introduced in 2020 would require waiting until approximately 2065 to observe complete CFR for even the oldest affected cohorts.

The CBR incorporates population composition and therefore captures the actual number of births occurring in a given year. This matters for fiscal planning, healthcare capacity, and education infrastructure, where the level of births rather than a synthetic rate is the relevant quantity. A policy could raise fertility rates while lowering the absolute number of births if it coincides with unusually small cohorts reaching childbearing age.

Working directly with cohort-by-age fertility responses resolves these limitations. Because TFR, CFR, and CBR are all aggregates of $\{f_{j,c}\}$, extrapolating the primitive responses allows construction of any aggregate measure. Given predicted responses $\{\widehat{\Delta f}_{j,c}\}$ for all treated cohort-age cells, one can simulate the path of TFR over time, recover completed CFR for cohorts still in their childbearing years, and forecast the trajectory of absolute births when combined with population projections.

The cohort-by-age responses are thus the common building block for all fertility measures. Extrapolating at this primitive level provides a unified framework rather than requiring

separate models for each aggregate. The decomposition developed in Section 3 imposes structure on these responses to enable extrapolation, while keeping the required assumptions explicit and testable.

2.3 The Extrapolation Problem

The fundamental challenge in fertility policy evaluation is that cohort-age responses are only partially observed at any given time. Cohorts fall into three groups based on their exposure to the policy and the extent to which their fertility histories are complete.

Untreated cohorts ($c + J_{\text{end}} < t_0$) complete childbearing before the policy is implemented. Their fertility is unaffected, so $\Delta f_{j,c} = 0$ for all ages j.

Partially treated cohorts ($c + J_{\text{start}} < t_0 \leq c + J_{\text{end}}$) are exposed to the policy at some fertile ages but not others. A woman born in 1950, for example, was 32 when the APFD began in 1982 and therefore experienced the policy only at ages 32 through 44. Responses are observed only for ages j satisfying $t_0 \leq c + j \leq t$, where t is the current observation date.

Fully treated cohorts ($c + J_{\text{start}} \geq t_0$) are exposed to the policy throughout their fertile years. For cohorts with $c + J_{\text{end}} > t$, however, responses at older ages have not yet occurred. A woman born in 1990 has been observed through age 35 by 2025 but will not complete childbearing until 2034.

At observation date t, the econometrician observes the set of realized responses:

$$\mathcal{R}^{\text{obs}}(t) = \big\{\Delta f_{j,c} : c + J_{\text{end}} \geq t_0,\ t_0 \leq c + j \leq t,\ j \in \{J_{\text{start}}, \ldots, J_{\text{end}}\}\big\}. \tag{5}$$

Evaluating completed fertility effects requires the full set of responses, including those not yet realized:

$$\mathcal{R}^{\text{full}} = \big\{\Delta f_{j,c} : c + J_{\text{end}} \geq t_0,\ c + j \geq t_0,\ j \in \{J_{\text{start}}, \ldots, J_{\text{end}}\}\big\}. \tag{6}$$

The gap between $\mathcal{R}^{\text{obs}}(t)$ and $\mathcal{R}^{\text{full}}$ defines the extrapolation problem: predicting fertility responses for cohort-age cells that have not yet been realized. Without additional structure, this gap is irreducible. There is simply no information in the data about fertility responses at ages cohorts have not yet reached.

The framework developed in Section 3 addresses this problem by imposing restrictions on the structure of cohort-age responses. The key assumption is that responses decompose additively into cohort and age components, with the age profile common across cohorts.

This structure, if valid, allows extrapolation of completed fertility from partial histories. The additive restriction is testable on observed cells, providing a diagnostic for when the extrapolation is warranted.

3 A Decomposition Framework

This section develops the decomposition framework that separates fertility responses into quantum and tempo components, enabling extrapolation of completed fertility effects before cohorts finish childbearing. The method proceeds in four steps: estimating cohort-by-age fertility responses, decomposing these responses into additive cohort and age fixed effects, extrapolating cohort effects for incompletely observed cohorts, and aggregating to construct policy-relevant measures. We then discuss the identifying assumptions underlying the decomposition and the extent to which they can be assessed empirically.

3.1 Step 1: Estimating Cohort-by-Age Responses

The first step recovers the primitive objects $\{\Delta f_{j,c}\}$ using reduced-form causal inference methods. We assume access to a control group—whether untreated units in a difference-in-differences design, a synthetic control, or a regression discontinuity comparison group—that provides a valid counterfactual for fertility in the absence of the policy.

For each cohort-age cell (c, j) observed in the post-policy period, the fertility response is estimated as the difference-in-differences coefficient comparing treated and control units. Specifically, for cohort bin c and age bin j:

$$\Delta \hat{f}_{j,c} = \left(\bar{f}_{j,c}^{\text{treated, post}} - \bar{f}_{j,c}^{\text{treated, pre}}\right) - \left(\bar{f}_{j,c}^{\text{control, post}} - \bar{f}_{j,c}^{\text{control, pre}}\right), \tag{7}$$

where the pre-period comparison at age j pools across cohorts observed at that age before the policy was implemented. This approach identifies the causal effect of the policy on fertility at each cohort-age cell under standard difference-in-differences assumptions: parallel trends in the absence of treatment and no anticipation of the policy.

The output of this step is a matrix of estimated responses $\{\Delta \hat{f}_{j,c}\}$ with associated standard errors, covering all cohort-age cells observed in the data. Cells requiring extrapolation—those where the cohort has not yet reached age j—remain missing at this stage. Section 5 implements this step for the APFD application, estimating separate treatment effects for each combination of 5-year birth cohort and 3-year age bin.

3.2 Step 2: Decomposing Responses into Cohort and Age Effects

The second step imposes an additive structure on the estimated responses. We regress the cell-level estimates on cohort and age fixed effects:

$$\Delta \hat{f}_{j,c} = \alpha_c + \beta_j + \varepsilon_{j,c}, \tag{8}$$

where α_c captures the cohort-specific fertility quantum (the total fertility effect for cohort c) and β_j captures the age-specific tempo pattern (how the policy effect is distributed across ages). We weight observations by the inverse of their estimated variance to account for heteroskedasticity in the cell-level estimates.

We normalize the age effects to sum to zero:

$$\sum_{j=J_{\text{start}}}^{J_{\text{end}}} \beta_j = 0. \tag{9}$$

This normalization serves two purposes. First, it resolves the standard location indeterminacy in two-way fixed effect models: without a normalization, one could add a constant to all α_c and subtract it from all β_j without changing the fit. Second, and more substantively, it ensures that the age effects represent pure timing shifts that do not affect completed fertility. Under this normalization, the cohort effect α_c directly measures the average per-age contribution to completed fertility for cohort c. The completed fertility rate response for cohort c is then:

$$\Delta \text{CFR}_c = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} (\alpha_c + \beta_j) \cdot \delta_j = \alpha_c \cdot \sum_{j=J_{\text{start}}}^{J_{\text{end}}} \delta_j, \tag{10}$$

where δ_j is the width of age bin j (e.g., 3 years for 3-year age bins). The second equality follows from the zero-sum normalization. The cohort fixed effect, scaled by the total years of childbearing, gives the completed fertility effect.

3.3 Step 3: Extrapolating Cohort Effects

The third step addresses the core extrapolation problem identified in Section 2.3: predicting cohort effects for cohorts not yet fully observed. For recent cohorts still in their childbearing years, responses are observed only at younger ages, yielding noisy estimates of the cohort effect. For future cohorts not yet in the data, nothing is observed directly. Extrapolation requires predicting the cohort effect based on the pattern observed for earlier cohorts.

The challenge here parallels the limited mobility problem in the AKM literature on wage decompositions. In that setting, worker fixed effects are poorly identified for workers observed at few firms, and variance component estimates can be severely biased when mobility is limited (Andrews et al., 2008; Bonhomme et al., 2019). The analog in the present context is that cohort effects are poorly identified for cohorts observed at few ages. A cohort observed only at ages 20–25 contributes limited information to separate its permanent fertility response from age-specific timing effects, just as a worker observed at a single firm contributes little to distinguishing worker ability from firm wage premia. The precision of cohort effect estimates increases mechanically with the number of ages over which the cohort is observed, creating a tradeoff between the relevance of recent cohorts and the reliability of their estimated effects.

We implement extrapolation using local polynomial regression, which fits a flexible trend through the observed cohort effects and projects forward. This approach balances bias and variance, capturing nonlinear patterns in the cohort effects without overfitting to noise in individual estimates. The predicted cohort effect for an unobserved or partially observed cohort c' is:

$$\hat{\alpha}_{c'} = \sum_{c \in \mathcal{C}^{\text{obs}}} w(c, c') \cdot \hat{\alpha}_c, \tag{11}$$

where $w(c, c')$ are kernel weights that decline with distance from c' and $\mathcal{C}^{\text{obs}}$ denotes the set of cohorts with sufficiently well-estimated effects.

The choice of weighting scheme involves a fundamental tradeoff between precision and relevance. More recent cohorts are more informative about current conditions, including general equilibrium effects of the policy that may have evolved over time, but are observed over fewer ages, yielding less precise cohort effect estimates. Earlier cohorts have more complete fertility histories and thus more precisely estimated cohort effects, but may be less relevant if the policy's effects have changed. The AKM literature offers guidance here: Kline et al. (2020) show that leave-out corrections can mitigate bias from including noisy fixed effect estimates in variance decompositions. In the extrapolation context, the analogous concern is that including imprecisely estimated cohort effects without appropriate downweighting can distort projected trends. We consider three weighting approaches that navigate this tradeoff.

Uniform weighting assigns equal weight to all observed cohorts within the kernel bandwidth, treating all cohort effects as equally informative for extrapolation. This approach maximizes the effective sample size but may be biased if cohort effects exhibit trends over time, and it fails to account for the heterogeneous precision of cohort effect estimates.

Precision weighting weights each cohort by the inverse variance of its estimated effect, $w_c \propto 1/\text{SE}(\hat{\alpha}_c)^2$. This approach downweights noisy estimates from cohorts observed over few ages, directly addressing the limited-observation problem. However, it may overweight older cohorts whose effects are precisely estimated yet potentially less relevant for current policy, introducing a form of staleness bias.

Recency weighting applies exponential decay weights, $w_c \propto \exp(-(c_{\max} - c)/\tau)$, where τ controls the half-life of relevance. This approach prioritizes recent cohorts under the assumption that their effects better reflect current policy impacts, at the cost of higher variance due to smaller effective sample sizes and greater influence of imprecisely estimated effects.

Our preferred specification combines precision and recency weighting, using $w_c \propto \exp(-(c_{\max} - c)/\tau)/\text{SE}(\hat{\alpha}_c)^2$. This hybrid approach prioritizes recent cohorts while downweighting those with imprecise estimates, balancing relevance against reliability. We use a Gaussian kernel with a bandwidth of 15 years, which provides sufficient smoothing while remaining responsive to trends in the cohort effects. For the main results in Section 5, we set $\tau = 20$ years, but we present results under all three individual weighting schemes to assess robustness to this choice.

For cohorts far into the future, all methods converge toward the local trend, and uncertainty grows correspondingly. We report extrapolations for five additional cohort bins (25 years forward), beyond which projections become increasingly speculative.

3.4 Step 4: Aggregating to Policy-Relevant Measures

The fourth and final step constructs aggregate fertility measures from the estimated and extrapolated cohort-age responses. With cohort effects $\{\hat{\alpha}_c\}$ (estimated for observed cohorts, extrapolated for future cohorts) and estimated age effects $\{\hat{\beta}_j\}$, we compute any aggregate that is a function of cohort-age-specific fertility.

The total fertility rate response in calendar year t sums age-specific effects across the cohorts present in that year:

$$\Delta\widehat{\text{TFR}}_t = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} (\hat{\alpha}_{t-j} + \hat{\beta}_j) \cdot \delta_j, \tag{12}$$

where δ_j is the width of age bin j and $t - j$ identifies the cohort of women at age j in year t. This measure captures how the policy affects the synthetic period fertility rate,

including both quantum and tempo components. Because different cohorts have different cohort effects, the TFR response varies over time as high-response and low-response cohorts move through the age distribution.

The completed fertility rate response for cohort c sums effects across all fertile ages, as given by equation (10):

$$\Delta\widehat{\text{CFR}}_c = \hat{\alpha}_c \cdot \sum_{j=J_{\text{start}}}^{J_{\text{end}}} \delta_j. \tag{13}$$

For cohorts that have completed childbearing, this represents actual completed fertility. For cohorts still in progress or not yet observed, this represents projected completed fertility under the extrapolated cohort effect.

Standard errors for these aggregates incorporate uncertainty from both the cell-level estimation and, where applicable, the extrapolation step. Under the assumption that estimation error is independent across cohort-age cells, the variance of the TFR response in year t is:

$$\text{Var}(\Delta\widehat{\text{TFR}}_t) = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} \left[\text{Var}(\hat{\alpha}_{t-j}) + \text{Var}(\hat{\beta}_j)\right] \cdot \delta_j^2. \tag{14}$$

For years relying on extrapolated cohort effects, the variance of $\hat{\alpha}_{t-j}$ incorporates uncertainty from the local polynomial regression.

This aggregation framework yields several outputs presented in Section 5: a time series of TFR responses showing how the policy affected period fertility year by year, a cohort series of CFR responses showing projected completed fertility by birth cohort, and decompositions of each into quantum and tempo components. Comparing TFR and CFR trajectories reveals whether observed period effects reflect permanent fertility changes or temporary timing shifts—the central question for long-run policy evaluation.

3.5 Discussion: Identification and Testability

The decomposition framework developed above rests on two key assumptions: the exclusion of period effects from the treatment effect structure, and the additive separability of cohort and age effects. This subsection discusses the relationship between these assumptions and the classical age-period-cohort identification problem, and clarifies what empirical tests can and cannot establish.

3.5.1 The Age-Period-Cohort Problem in Treatment Effects

The decomposition in equation (8) deliberately excludes period (calendar year) effects, which raises the question of how this framework relates to the classical age-period-cohort (APC) identification problem. In standard APC analysis of fertility levels, the accounting identity $t = c + j$ creates perfect collinearity among age, period, and cohort, making it impossible to separately identify all three effects without imposing additional restrictions (Glenn, 2005; Yang and Land, 2013).

One might initially suppose that this collinearity problem does not apply to treatment effects, since the difference-in-differences design removes period shocks common to treated and control groups. However, the APC problem re-emerges if treatment effects themselves have period-specific components. Consider the more general model:

$$\Delta f_{j,c} = \alpha_c + \beta_j + \gamma_t, \tag{15}$$

where γ_t captures period-specific variation in treatment effects and $t = c+j$. The collinearity remains: for any scalar λ, the transformation $\alpha_c^* = \alpha_c + \lambda c$, $\beta_j^* = \beta_j + \lambda j$, and $\gamma_t^* = \gamma_t - \lambda t$ leaves the sum $\alpha_c + \beta_j + \gamma_t$ unchanged. This indeterminacy cannot be resolved by normalizations alone—the zero-sum constraint $\sum_j \beta_j = 0$ that we impose pins down the level of age effects relative to cohort effects, but does not eliminate the fundamental identification problem when period effects are present.

Our baseline framework therefore relies on a substantive assumption:

Assumption 1 (No Period Effects in Treatment Effects)**.** *After removing common period shocks via the difference-in-differences design, the remaining treatment effects contain no period-specific component:* $\gamma_t = 0$ *for all* t *in equation* (15)*.*

This assumption would be violated if the policy's effectiveness varied with calendar time in ways specific to the treated group. Several channels could generate such period effects. First, if the policy's intensity varies over time—for instance, if the real value of the Alaska dividend fluctuates with inflation or investment returns—the treatment "dose" varies by calendar year, generating period-specific effects. Second, if contemporaneous economic conditions interact with the policy in ways specific to the treated group, such as oil price shocks that affect Alaska's economy differently than control states, the treatment effect could vary by period even after the difference-in-differences adjustment. Third, if the composition of the treated

population changes over time through selective migration, the average treatment effect at each period would reflect compositional shifts rather than changes in individual responses.

We view Assumption 1 as reasonable in our application for two reasons. First, the difference-in-differences design absorbs period-level variation that is common across treated and control groups. What remains in $\Delta\hat{f}_{j,c}$ is the differential response of Alaska relative to other states, which should not contain period effects unless Alaska-specific shocks interact with the policy. Second, the conceptual structure of fertility responses supports a cohort-age rather than period framework: a woman's fertility decision at age j depends on her cumulative policy exposure (cohort), her position in the life cycle (age), and idiosyncratic factors, but less plausibly on the calendar year per se after conditioning on cohort and age.

We acknowledge that Assumption 1 is fundamentally untestable in the strict sense—any test for period effects faces the same identification problem that motivates the assumption. However, we can assess the sensitivity of our conclusions by estimating an alternative specification that allows for period effects under a different identifying restriction. Appendix B implements the approach of Heckman et al. (1998), which identifies period effects by assuming that age effects are zero for the oldest age groups in our sample (ages 39–41 and 42–44). This assumption is motivated by the biological reality that fertility responses to income shocks are minimal near the end of the reproductive span. The results provide reassurance on two fronts. First, the estimated period effects are small in magnitude—fluctuating between approximately -0.005 and $+0.005$ additional births per woman per age-year—and statistically indistinguishable from zero throughout most of the sample period. Second, the cohort effects from this alternative specification closely track those from our baseline approach, with both series exhibiting the same qualitative pattern and the baseline estimates falling within or near the alternative specification's confidence intervals for cohorts born between 1940 and 2005. The primary divergence occurs at the boundaries of the cohort distribution, where identification is inherently more fragile. This robustness across identification strategies suggests that our substantive conclusions do not hinge critically on the no-period-effects assumption.

3.5.2 Testable Implications of Additive Separability

Beyond the exclusion of period effects, equation (8) embodies a second key assumption: the timing of fertility responses (the age profile $\{\beta_j\}$) is common across cohorts, while only the total magnitude of the response (the quantum α_c) varies by cohort. This additive separability

assumption is motivated by a model of intertemporal fertility choice in which the policy affects the timing of births in a uniform way—for example, by relaxing liquidity constraints that bind at young ages—while the total fertility response depends on cumulative exposure and cohort-specific characteristics.

Unlike Assumption 1, the additive separability of cohort and age effects is partially testable on observed cells. We assess model fit by examining the residuals $\varepsilon_{j,c}$ and testing for systematic departures from additivity. A linear interaction term $\gamma \cdot (c - \bar{c})(j - \bar{j})$ tests whether the age gradient of effects varies systematically across cohorts:

$$\Delta \hat{f}_{j,c} = \alpha_c + \beta_j + \gamma \cdot (c - \bar{c})(j - \bar{j}) + \varepsilon_{j,c}. \tag{16}$$

The interpretation of this test requires care given the APC structure. Since $t = c + j$, the interaction $(c - \bar{c})(j - \bar{j})$ is correlated with variation along the diagonals of the cohort-age matrix—precisely where period effects would manifest. A significant γ could therefore indicate either cohort-specific age profiles (failure of additive separability), period effects in treatment effects (failure of Assumption 1), or some combination of both. The test cannot distinguish between these possibilities.

However, failure to reject $\gamma = 0$ provides meaningful reassurance. If the data were generated by a model with substantial period effects or cohort-specific timing patterns, we would expect to detect this as a significant interaction, at least when effects are linear in the relevant dimension. Thus, while the test is not definitive, it serves as a useful diagnostic: rejection would cast doubt on the maintained assumptions and suggest that extrapolation requires either a richer model or stronger identifying restrictions.

Section 5.7 implements this test for the APFD application and finds no evidence against additivity ($p = 0.67$), with the two-way decomposition explaining nearly half the variation in cell-level treatment effects ($R^2 = 0.47$). We also examine residual patterns across the cohort-age surface for systematic structure that the linear interaction test might miss. The absence of such patterns, combined with the near-zero period effects and robust cohort effects documented in Appendix B, provides convergent support for the maintained specification. We emphasize, however, that these diagnostics cannot definitively rule out all violations of our identifying assumptions—they establish only that any such violations are sufficiently small or sufficiently structured that they do not substantially affect our conclusions.

4 Comparison with Alternative Approaches

Estimating long-run fertility effects from short-run data is a pervasive challenge in policy evaluation. This section situates the present framework within the landscape of alternatives. A key observation is that data requirements across approaches are essentially equivalent: any researcher with age-specific fertility rates by cohort for treated and control groups can implement the decomposition framework developed in Section 3. The differences lie not in data but in the structure imposed and the questions each approach is best suited to answer.

Waiting for Completed Fertility

The most conservative approach is to refrain from extrapolation entirely and wait until treated cohorts complete childbearing. This is methodologically unimpeachable, requiring no assumptions about future behavior, but often impractical. Cohorts fully exposed to a policy introduced in 1982 do not complete fertility until the mid-2020s, by which point the policy environment may have changed substantially and the original evaluation question may have lost relevance. For more recent policies, the wait is correspondingly longer. Policymakers deciding whether to expand, modify, or terminate a program cannot always defer judgment for decades.

Period Total Fertility Rate

The dominant approach in applied work is to estimate effects on period total fertility rates. This requires no assumptions beyond those needed for causal identification, making it a natural starting point for policy evaluation. The limitation is interpretive rather than methodological. A policy that accelerates childbearing raises period TFR mechanically even if completed fertility is unchanged, while a policy that raises completed fertility but induces delay can show muted or even negative period effects in the short run. Period TFR answers a well-defined question, namely how did births in a given year change, but this may differ fundamentally from the question of ultimate policy interest: how did the number of children women have over their lifetimes change?

The empirical analysis in Section 5 illustrates this divergence. The APFD raised period TFR substantially at its peak, but this figure overstates the effect on completed fertility because it includes tempo effects that sum to zero over the life course.

Tempo-Adjusted TFR

Bongaarts and Feeney (1998) proposed adjusting period TFR for tempo distortions using the formula $\text{TFR}^* = \text{TFR}/(1-r)$, where r is the annual change in mean age at childbearing. This approach has been influential in formal demography and offers a valuable correction when timing shifts would otherwise distort period measures.

The two frameworks impose different structural assumptions. The Bongaarts-Feeney adjustment assumes tempo shifts are uniform across ages and parities—all women delay or accelerate childbearing by the same amount regardless of where they are in the fertility distribution. The present framework instead assumes timing shifts are uniform across cohorts: the age profile of policy responses is common, while total magnitude varies by cohort. Neither assumption is strictly weaker. Bongaarts-Feeney uniformity may fail when policies differentially affect first versus higher-order births; cohort-uniformity may fail if the policy's timing effects evolve as social norms adapt.

A key practical difference is integration with causal inference. The Bongaarts-Feeney adjustment requires only aggregate period data, making it difficult to embed in a difference-in-differences design—one cannot easily construct a "tempo-adjusted treatment effect" because the adjustment is nonlinear and the treated group's tempo distortion reflects both policy effects and underlying trends. The cohort-age decomposition operates on the same cell-level treatment effects that standard policy evaluation produces, inheriting the identification properties of the underlying research design.

The approaches also answer different questions. Bongaarts-Feeney asks: what would period TFR be if current timing shifts continued indefinitely? The present framework asks: what is the completed fertility effect for a specific cohort exposed to a specific policy, relative to an explicit counterfactual? The Bongaarts-Feeney adjustment is thus well suited for characterizing aggregate national trends where no control group exists (Sobotka, 2004; Goldstein et al., 2009), while the cohort-age decomposition is designed for policy evaluation settings where a credible counterfactual is available and the goal is to assess whether a specific intervention changed completed fertility.

Mechanical Projection

A related class of methods imputes remaining fertility for incomplete cohorts by assuming they will follow the age-specific rates of some reference population, typically an earlier cohort or a control group, or that recent trends will continue. These methods are transparent and

easy to implement. The challenge is that the projection assumptions are difficult to evaluate: there is no formal basis for assessing whether the reference population is appropriate or whether trend extrapolation is warranted in a particular context.

The present framework can be viewed as a structured version of mechanical projection, where the projection assumptions are made explicit through the additive decomposition and can be partially tested on observed data. Rather than assuming that a recent cohort's age-specific rates will continue unchanged, the framework estimates a common age profile from all observed cohorts and extrapolates only the cohort effects.

Structural Life-Cycle Models

At the opposite extreme, structural life-cycle models explicitly specify preferences, constraints, and dynamic optimization over the fertility career (Keane and Wolpin, 1997; Francesconi, 2002; Gayle and Miller, 2012). These models can extrapolate to completed fertility by forward simulation under alternative policy scenarios and provide rich economic interpretation of the mechanisms through which policies operate.

The costs of the structural approach are substantial. Identification requires functional form assumptions on utility and budget constraints, data demands often extend to panel data on individual fertility and labor market histories, and computational complexity limits robustness analysis. The structural approach is most valuable when understanding mechanisms is the primary goal and the required data are available. The present framework prioritizes transparency and testability for policy evaluation, sacrificing some interpretive richness for empirical tractability.

The Present Framework

The decomposition framework developed in Section 3 combines strengths of these alternatives while avoiding their main limitations. It shares with reduced-form approaches the reliance on transparent, testable assumptions applied to standard repeated cross-section data. It shares with tempo-adjustment methods the goal of separating timing shifts from completed fertility changes, though it implements this separation through cohort and age fixed effects rather than through adjustments based on the rate of change in mean age.

The key maintained assumptions are that fertility responses decompose additively into cohort and age components (with no residual period effects, per Assumption 1), that the age profile of responses is common across cohorts, and that cohort effects evolve smoothly

over time. These assumptions are partially testable on observed data. The additive structure can be assessed by examining residuals and testing for interactions (as implemented in Section 5.7), while the smoothness of cohort effects can be evaluated visually and through specification tests.

Tradeoffs and Limitations

Each approach embodies a different tradeoff between robustness and informativeness. Period TFR makes minimal assumptions but cannot distinguish timing from quantum. Tempo-adjusted TFR corrects for aggregate timing shifts but requires uniformity assumptions and is designed for population description rather than policy evaluation. Mechanical projection is flexible but lacks a framework for evaluating its assumptions. Structural models are rich but require strong functional form commitments. The present approach offers a disciplined middle path, imposing enough structure to separate quantum from tempo and to extrapolate completed fertility while keeping those assumptions explicit and empirically testable.

This middle path comes with its own limitations. Like the other descriptive approaches discussed above, the framework measures how much completed fertility changes but does not identify the underlying mechanisms. Whether effects operate through income effects, substitution effects, changes in marriage markets, or other channels remains unresolved. The extrapolation of cohort effects to future cohorts is inherently uncertain. There is no guarantee that future cohorts will follow the trends estimated from past cohorts, and the further forward the projection extends, the greater this uncertainty becomes.

Assumption 1—that treatment effects contain no period-specific component after the DID adjustment—is imposed rather than derived, and if this assumption fails, the decomposition will be misspecified. The same applies to the assumption that the age profile of responses is constant across cohorts. Finally, the framework requires a credible control group or counterfactual to identify the cohort-age-specific treatment effects in the first step. The decomposition and extrapolation methods impose structure on estimated effects but do not solve the fundamental identification problem.

Despite these limitations, for the common problem of evaluating completed fertility effects before cohorts finish childbearing, the present framework offers clear advantages. It extracts more information from the same data that researchers already use for period TFR analysis: separate estimates of quantum and tempo effects, extrapolation of completed fertility under transparent and testable assumptions, and uncertainty quantification that reflects both esti-

mation error and projection uncertainty. The assumptions are stronger than those required for period TFR alone, but they are explicit rather than hidden and can be evaluated against the data. When the goal is to assess long-run fertility impacts of policy without waiting decades for completed fertility data, this framework provides a direct path from standard difference-in-differences estimates to the answer policymakers need.

5 Application: The Alaska Permanent Fund Dividend

This section applies the decomposition framework to the Alaska Permanent Fund Dividend (APFD), demonstrating the method in practice and illustrating the importance of the quantum-tempo distinction. We first describe the institutional setting and data, then present estimates of cohort-by-age treatment effects, and finally implement the four-step decomposition to recover tempo effects, quantum effects, and extrapolated completed fertility.

5.1 Institutional Background

The Alaska Permanent Fund was established in 1976 as a constitutional amendment requiring that at least 25 percent of the state's oil royalties be deposited into a dedicated investment fund. Beginning in 1982, the state began distributing annual dividends to Alaska residents from the fund's investment earnings. The Alaska Permanent Fund Dividend (APFD) is paid to all residents who have lived in the state for at least one full calendar year, with no restrictions based on age, employment status, or family composition.

The dividend amount varies from year to year based on investment performance, ranging from approximately $300 to over $2,000 per person in nominal terms. In 2024, the dividend was $1,702 per person, meaning a family of four received $6,808. Over the program's history, cumulative payments to a long-term resident have exceeded $50,000. The APFD thus represents a substantial, recurring, and unconditional cash transfer, making it a valuable setting for studying fertility responses to income.

One consideration for identification is whether the APFD induces selective migration. The one-year residency requirement could theoretically attract families planning to have children or retain families who might otherwise leave the state. Yonzan et al. (2024) note that fertility trends in Alaska during the late 1970s were influenced by migration associated with oil pipeline construction, though they treat this as a pre-treatment confounder addressed by synthetic control matching rather than a policy-induced response. Their analysis does not

find evidence that the APFD itself generated differential migration, and Jones and Marinescu (2022) similarly find no significant labor market distortions from the dividend. We follow the existing literature in treating migration responses as second-order, while acknowledging that any migration-induced compositional changes would affect both the period estimates from which we begin and the cohort effects we recover.

5.2 Data

The empirical analysis uses micro-level data from the Current Population Survey (CPS), a nationally representative household survey conducted monthly by the U.S. Census Bureau. We aggregate monthly observations to the annual level and construct a sample spanning 1976 to 2025, comprising a pre-period of six years (1976 to 1981) before the APFD was implemented and a post-period of over four decades (1982 to 2025).

The sample is restricted to women of childbearing age, defined as ages 15 to 44. A birth in the current year is identified using household composition variables: a woman is coded as having a new birth if she reports having children in the household and her youngest child is less than one year old. This approach is standard in CPS-based fertility research (Lindo, 2010), though it may miss births where the child is not present in the household at the time of the survey.

The full sample comprises over 17 million person-year observations of women of childbearing age, of which approximately 250,000 are Alaska residents. Women in the remaining U.S. states serve as the control group in the difference-in-differences framework. Age-specific fertility rates are computed by dividing the number of women with a new birth by the total number of women at each age, separately for Alaska and the control states. All tabulations incorporate CPS survey weights.

The extended timeframe is crucial for the decomposition exercise. By 2025, cohorts who entered childbearing in the early 1980s have largely completed their fertility, providing a benchmark against which to assess the extrapolation methods. At the same time, cohorts born in the 1990s and later are still in their childbearing years, providing the variation needed to test whether the additive structure holds and to evaluate extrapolation uncertainty.

5.3 Estimation Strategy

Following the framework developed in Section 3, estimation proceeds in three steps: estimating cohort-by-age specific treatment effects using difference-in-differences, decomposing

these effects into cohort and age components, and extrapolating cohort effects to project completed fertility for cohorts still in their childbearing years.

Cohort-Age Specific Treatment Effects

For each combination of 5-year birth cohort bin c and 3-year age bin a, we estimate a difference-in-differences regression:

$$Y_{ist} = \beta_{ca} \cdot (\text{Alaska}_s \times \text{Post}_t) + \gamma_t + \delta_s + \varepsilon_{ist}, \tag{17}$$

where Y_{ist} is an indicator for whether woman i in state s and year t had a birth, Alaska_s indicates Alaska residence, Post_t indicates the post-1982 period, and γ_t and δ_s are year and state fixed effects. The coefficient β_{ca} captures the differential change in fertility for Alaska women in cohort c at age a, relative to women in other states.

The pre-period (1976 to 1981) establishes the baseline difference between Alaska and other states at each age, pooling across all cohorts observed in those years. The post-period comparison is specific to each cohort-age cell: for cohort c at age a, the post-period includes only years when that cohort was actually age a (years $c + a$ through $c + a + 2$ for a 3-year age bin). Standard errors are heteroskedasticity-robust.

This procedure yields a matrix of treatment effects $\{\hat{\beta}_{ca}\}$ across 17 cohort bins (midpoints from 1940 through 2020 in 5-year intervals) and 10 age bins (15 to 17, 18 to 20, and so on through 42 to 44), though not all cells are populated due to the structure of the data. Older cohorts are observed only at older ages in the post-period; younger cohorts have not yet reached older ages by 2025.

A key identifying assumption underlying this difference-in-differences approach is that fertility trends in Alaska would have paralleled those in other states absent the APFD. While this parallel trends assumption cannot be directly tested, we provide supporting evidence in Appendix A. We implement a synthetic control approach that constructs a data-driven counterfactual for Alaska by matching on pre-treatment fertility and demographics, providing a robustness check that relaxes the equal-weighting implicit in standard DID.

Decomposition into Quantum and Tempo

We decompose the estimated treatment effects into cohort and age components by estimating equation (8):

$$\hat{\beta}_{ca} = \mu + \alpha_c + \beta_a + \eta_{ca}, \tag{18}$$

where α_c represents the cohort fixed effect (quantum) and β_a represents the age fixed effect (tempo). The regression is estimated by weighted least squares, with weights equal to the inverse variance of the first-stage estimates ($1/\hat{\sigma}^2_{ca}$), giving more influence to more precisely estimated cells.

The age fixed effects are normalized to sum to zero, $\sum_a \beta_a = 0$, so that they capture pure timing shifts that do not affect completed fertility. The cohort fixed effects are identified relative to this normalization, so that α_c directly measures the per-age contribution to completed fertility for cohort c attributable to the policy.

Extrapolation of Cohort Effects

Cohorts born after approximately 1980 have not yet completed their fertile years by 2025, so their cohort fixed effects reflect only partially observed fertility histories. To project completed fertility for these cohorts, we smooth and extrapolate the cohort effects using local polynomial regression with a Gaussian kernel (bandwidth of 15 years), as described in Section 3.

We implement three weighting schemes: uniform weighting, precision weighting by inverse variance of the cohort effects, and recency weighting with exponential decay (half-life of 20 years). The extrapolation extends 25 years (five 5-year cohort bins) beyond the last observed cohort. For the main results, we use recency weighting, which balances responsiveness to recent trends against stability from the broader sample.

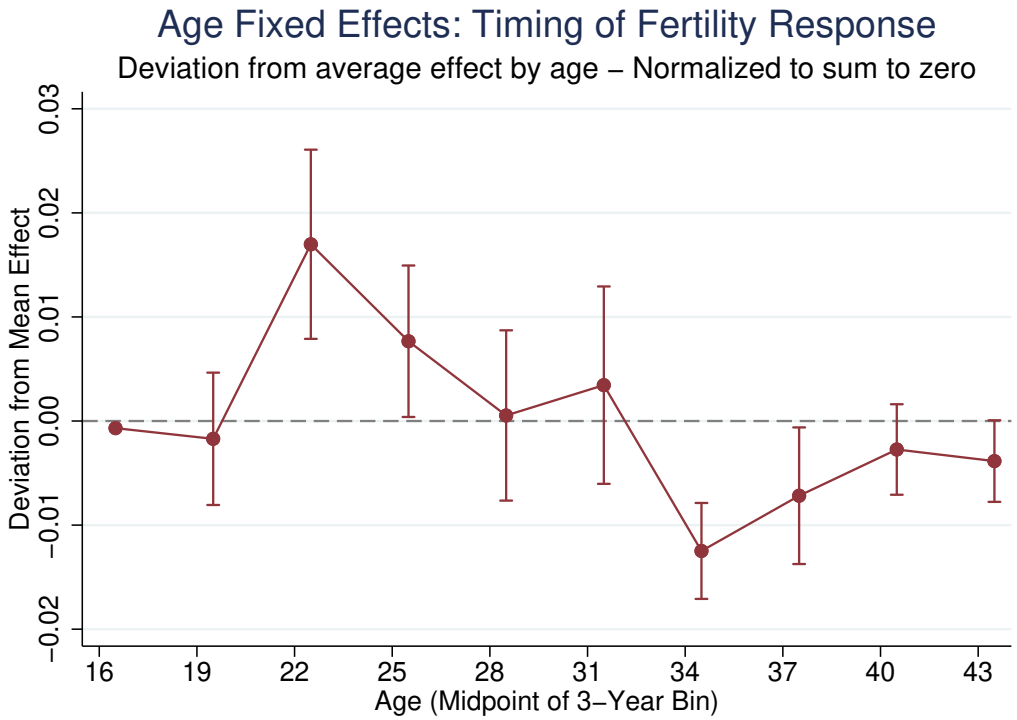
5.4 Results: Age Effects (Tempo)

Figure 1 presents the estimated age fixed effects, which capture how the APFD shifted the timing of fertility within cohorts. The pattern shows elevated fertility at younger ages, peaking around ages 21 to 23, and reduced fertility at older ages, most negative around ages 33 to 35. This profile indicates that the APFD accelerated childbearing: women had children earlier than they otherwise would have.

The magnitude of the tempo effect is substantial. At its peak, the APFD increased

the probability of a birth by approximately 1.7 percentage points for women in their early twenties, relative to the average effect across all ages. The negative effects at older ages are smaller in absolute magnitude but more precisely estimated, as these cells contain larger samples and longer observation windows.

Because the age effects sum to zero by construction, these timing shifts do not contribute to changes in completed fertility. A woman who has a child at age 22 instead of age 32 due to the APFD contributes to the positive age effect at young ages and the negative effect at older ages, but her total number of children is unchanged. The tempo effects documented in Figure 1 are substantial in their short-run impact on period fertility but transitory in their implications for family size.



Notes: Points show estimated age fixed effects from the decomposition regression (equation 8), normalized to sum to zero across ages. Whiskers indicate 95% confidence intervals. Positive values at younger ages and negative values at older ages indicate that the APFD accelerated the timing of childbearing. The synthetic control analysis in Appendix A recovers a nearly identical age profile (Figure 6), with a correlation exceeding 0.9 between DID and synthetic control estimates.

Figure 1. Estimated Age Fixed Effects (Tempo)

5.5 Results: Cohort Effects (Quantum)

Figure 2 presents the estimated cohort fixed effects, which measure the total fertility response to the APFD for each birth cohort. These effects, when scaled by the 30 years of childbearing (10 age bins multiplied by 3 years per bin), yield the completed fertility rate response.

Several patterns emerge from the cohort effects. First, cohorts born in the late 1960s, who were young adults when the APFD began in 1982 and experienced the policy throughout most of their childbearing years, show the largest positive effects. Second, earlier cohorts (born in the 1940s and 1950s) show smaller but still positive effects. These cohorts were already partway through their childbearing years when the APFD began, so they experienced less cumulative exposure to the policy. The 1940 cohort, for example, was already 42 years old in 1982 and thus had only two to three remaining years of potential fertility.

Third, more recent cohorts (born in the 1980s and later) show effects closer to zero or slightly negative. These estimates are less reliable because these cohorts have not yet completed childbearing, so their cohort effects are identified from fewer observed ages. The apparent decline in cohort effects for recent cohorts may reflect genuine changes in policy response, changes in the composition of Alaska's population, or simply noise from incomplete observation.

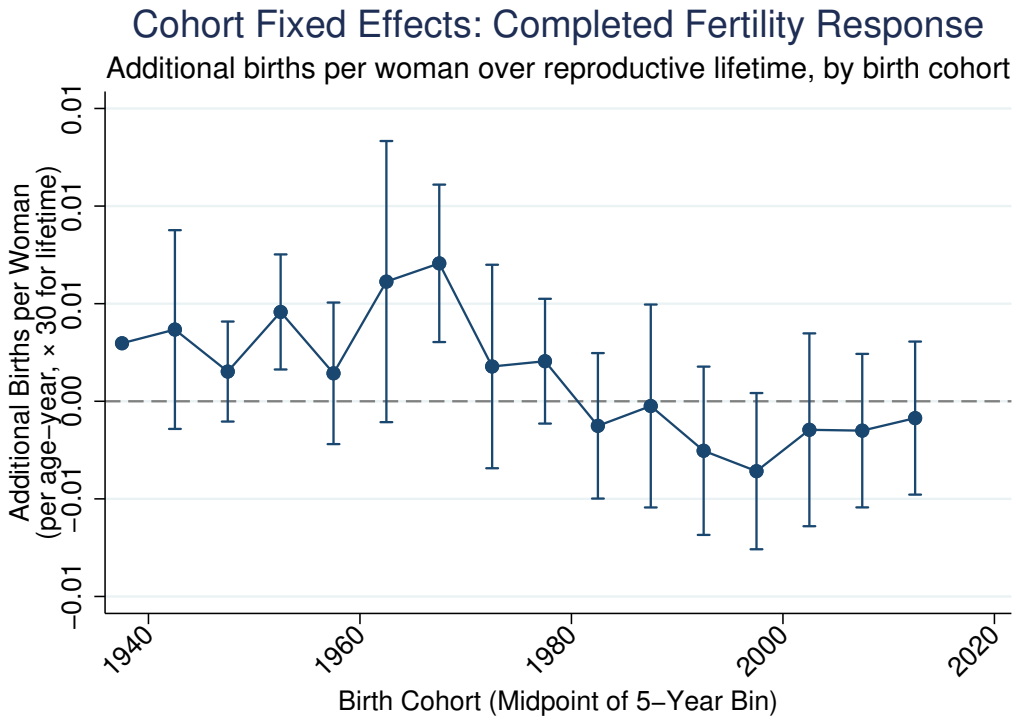
5.6 Extrapolation and Projections

Extrapolated Cohort Effects

Figure 3 presents the extrapolation of cohort effects to partially observed and future cohorts. Solid points show the estimated cohort fixed effects from the decomposition regression, with 95% confidence intervals. The three curves show smoothed fits from local polynomial regression under the three weighting schemes described in Section 3. Hollow points to the right of the vertical dashed line show extrapolated values for five additional cohort bins (cohorts born 2020 to 2024 through 2040 to 2044), using each of the three methods.

Several patterns emerge from the extrapolation. First, all three smoothing methods capture the general shape of the observed cohort effects, with effects peaking for cohorts born in the late 1960s and declining for more recent cohorts. The methods diverge primarily in their treatment of the most recent observed cohorts, which have noisier estimates due to incomplete fertility histories.

Second, the precision-weighted method produces a flatter trajectory, as it downweights re-



Notes: Points show estimated cohort fixed effects from the decomposition regression (equation 8). Whiskers indicate 95% confidence intervals. The cohort effect measures the per-age contribution to completed fertility; multiplying by 30 (the number of childbearing years) gives the total completed fertility rate response. Cohorts born in the late 1960s show the largest effects, consistent with full exposure to the APFD during their prime childbearing years. The synthetic control estimates in Appendix A show a similar pattern, though with greater volatility for recent cohorts (Figure 7).

Figure 2. Estimated Cohort Fixed Effects (Quantum)

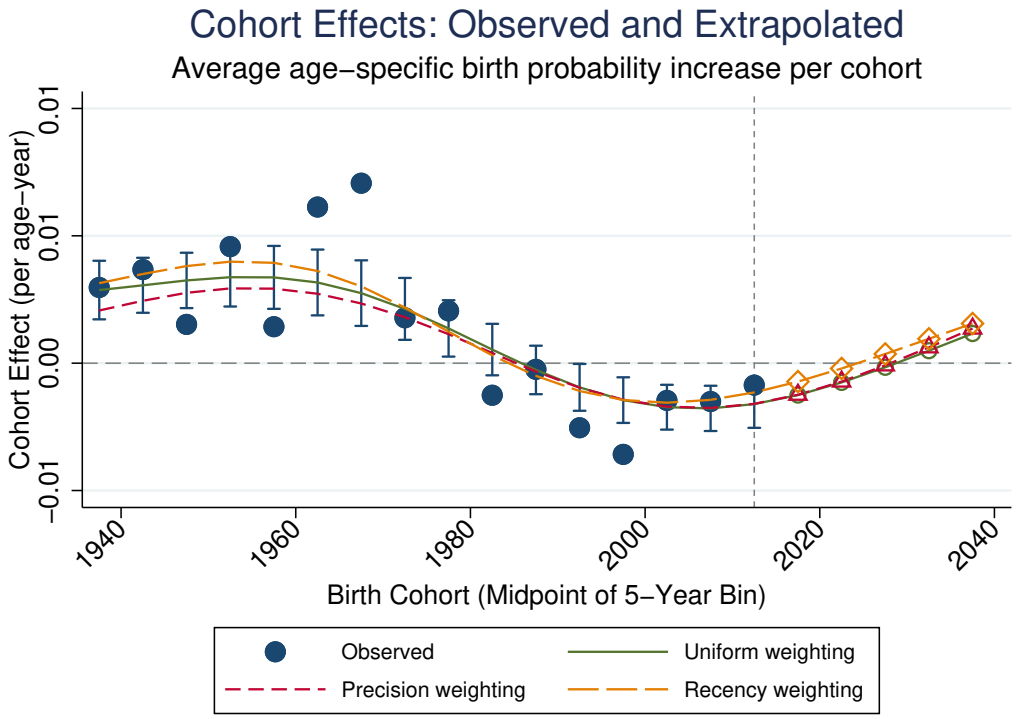
cent cohorts whose estimates have larger standard errors. The uniform and recency-weighted methods track each other closely through the observed period but diverge slightly in the extrapolation region.

Third, the extrapolated cohort effects for future cohorts (born in the 2020s through 2040s) are close to zero under all methods, indicating that the APFD's effect on completed fertility has largely dissipated for cohorts entering childbearing in recent years.

Completed Fertility Rate Response

Figure 4 presents the completed fertility rate (CFR) response by birth cohort, calculated as the cohort fixed effect multiplied by 30 years of childbearing. The CFR response is expressed as additional births per woman over the reproductive lifespan.

Cohorts born in the late 1960s experienced the largest fertility response. Earlier co-

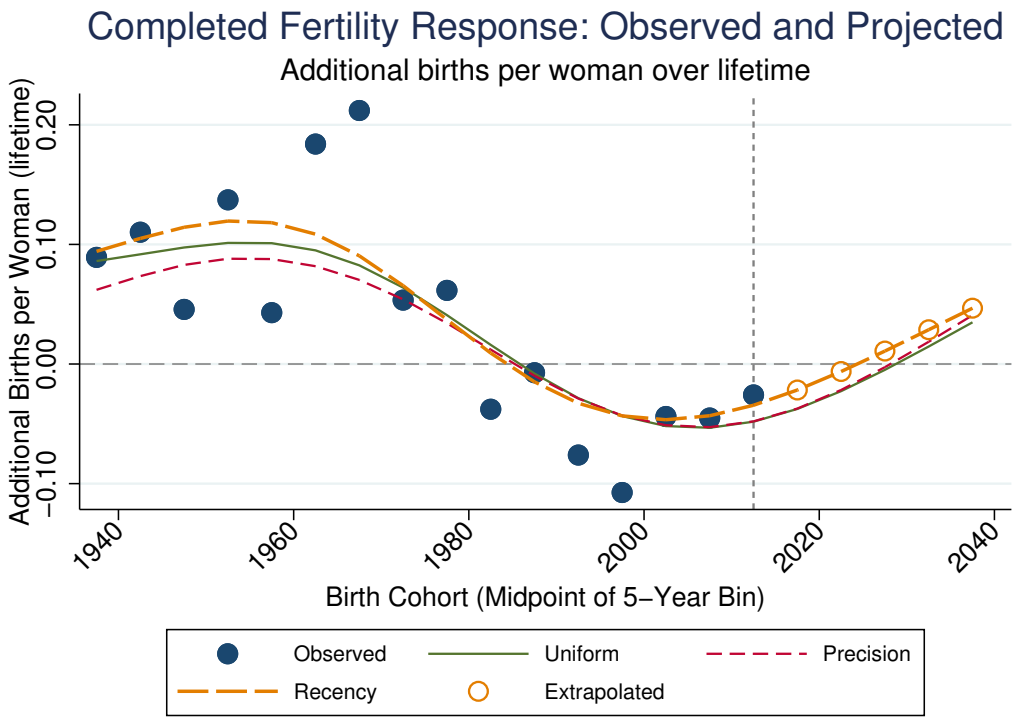


Notes: Solid points show estimated cohort fixed effects with 95% confidence intervals. Lines show smoothed fits from local polynomial regression under three weighting schemes: uniform (solid), precision (short-dashed), and recency (long-dashed). Hollow points show extrapolated values using each method. The vertical dashed line separates cohorts with complete or nearly complete fertility histories (left) from those requiring substantial extrapolation (right).

Figure 3. Cohort Fixed Effects: Observed and Extrapolated

horts show smaller positive effects, consistent with their partial exposure to the policy. For cohorts born after 1980, the observed CFR effects decline toward zero, and the smoothed extrapolation suggests these cohorts will have negligible completed fertility effects when their childbearing is complete.

The extrapolated CFR effects for future cohorts (born in the 2020s through 2040s) cluster tightly around zero, indicating that the APFD's effect on completed fertility has effectively ended for cohorts not yet in the data. This finding has important implications for long-run population projections: the APFD's fertility effects are largely a transitional phenomenon affecting cohorts who experienced the policy during their prime childbearing years, rather than a permanent feature of Alaska's demographic regime.



Notes: CFR response equals the cohort fixed effect multiplied by 30 years of childbearing, expressed as additional births per woman over the lifetime. Solid points show observed cohorts; hollow points show extrapolated values. Lines show smoothed projections under three weighting schemes: uniform, precision, and recency. The peak CFR response occurs for cohorts born in the late 1960s.

Figure 4. Completed Fertility Rate Response by Cohort

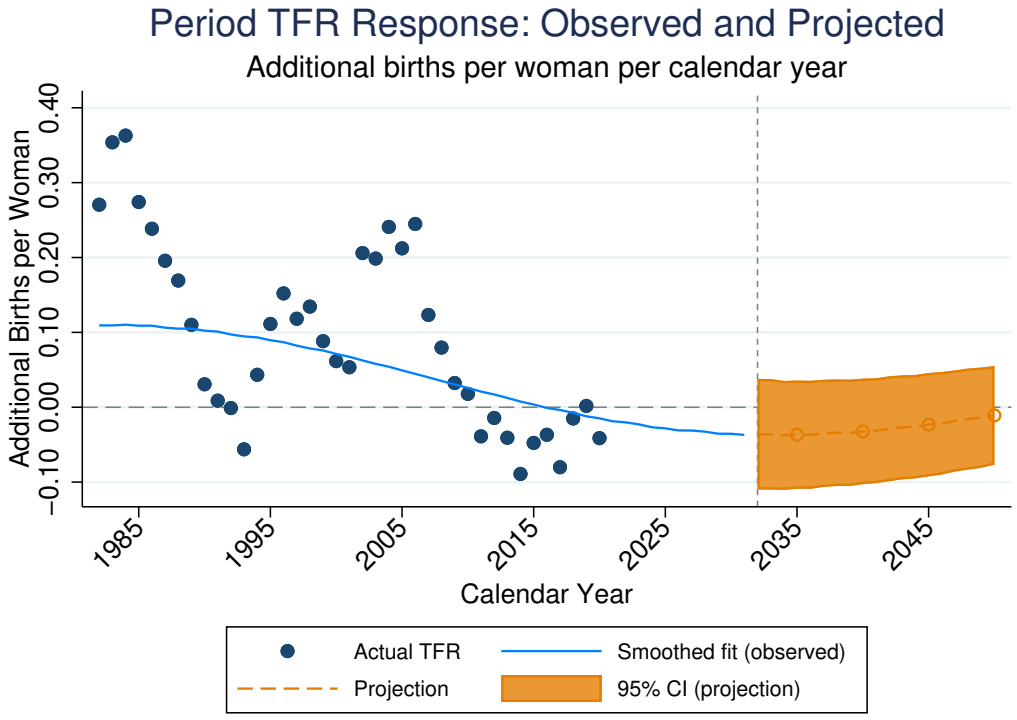
Total Fertility Rate Response

Figure 5 presents the TFR response by calendar year, constructed by aggregating cohort and age effects as described in equation (12). The TFR response measures how the APFD affected period fertility in each year, combining both quantum effects (changes in completed fertility) and tempo effects (shifts in timing).

The TFR response shows a pronounced declining pattern over time. In the early 1980s, immediately following the introduction of the APFD, the TFR response was at its highest, with substantial additional births per woman attributable to the policy. By the 2010s, the TFR response had declined substantially, and projections through 2050 show it remaining close to zero.

The declining TFR trajectory reflects two distinct forces. First, the cohorts with the largest completed fertility responses (born in the 1960s) have aged out of childbearing, re-

placed by younger cohorts with smaller estimated quantum effects. Second, the tempo effects documented in Figure 1, the acceleration of childbearing to younger ages, generate a temporary boost to the TFR that dissipates as the timing shift works through the population. A woman who has her children earlier due to the APFD contributes to elevated TFR in those earlier years but correspondingly lower TFR in later years when she would otherwise have given birth.



Notes: Points show the TFR response calculated from observed cohort-age cells. The solid line shows a smoothed fit through the observed period; the dashed line and shaded region show projections with 95% confidence intervals for years using extrapolated cohort effects. The peak TFR response in the mid-1980s substantially exceeds the peak CFR response, reflecting tempo effects that inflated period fertility.

Figure 5. Total Fertility Rate Response by Calendar Year

Interpreting the Quantum-Tempo Gap

The contrast between the TFR and CFR trajectories quantifies the importance of the quantum-tempo distinction emphasized throughout this paper. The peak TFR response substantially exceeds the peak CFR response. This gap reflects pure tempo effects: births that were shifted earlier in time but would have occurred regardless.

The age fixed effects in Figure 1 provide the mechanism behind this gap. Women in their early twenties showed fertility increases relative to the mean, while women in their mid-thirties showed reductions. Because the age effects sum to zero by construction, these timing shifts cancel out over the life course and do not contribute to completed fertility. They do, however, contribute substantially to the period TFR in early years, generating the wedge between TFR and CFR trajectories.

For policymakers, this decomposition carries an important lesson. A naive interpretation of the 1980s TFR effects would conclude that the APFD substantially increased Alaska's birth rates. The decomposition reveals that much of this increase was illusory, a timing shift that borrowed births from the future rather than creating new ones. The true effect on completed fertility, while still positive for some cohorts, was smaller than the period TFR suggested.

5.7 Specification Tests and Robustness

The decomposition framework relies on two maintained assumptions: Assumption 1 (no period effects in treatment effects) and additive separability of cohort and age effects. This subsection evaluates the empirical support for these assumptions.

We test for violations of additivity by estimating the augmented regression from equation (16):

$$\Delta \hat{f}_{j,c} = \alpha_c + \beta_j + \gamma \cdot (c - \bar{c})(j - \bar{j}) + \varepsilon_{j,c}, \tag{19}$$

where $(c - \bar{c})$ and $(j - \bar{j})$ are cohort and age deviations from their weighted means, and γ captures whether the age gradient of effects varies systematically across cohorts. As discussed in Section 3, a significant γ could indicate either cohort-specific age profiles or period effects in treatment effects; the test cannot distinguish between these possibilities.

Table 1 reports the results. The estimated interaction coefficient is $\hat{\gamma} = 0.000005$, with a standard error of 0.000012, yielding a t-statistic of 0.43 and a p-value of 0.67. The interaction term is economically negligible and statistically insignificant, providing no evidence against the additive specification.

This finding supports the maintained assumptions. The APFD appears to have shifted the timing of fertility in a similar way for all cohorts, accelerating childbearing toward younger ages, with only the overall magnitude of the response varying by cohort. This stability of the age profile validates the extrapolation exercise: projecting cohort effects

	Interaction Term
Coefficient	0.000005
Standard Error	0.000012
t-statistic	0.431
p-value	0.668
Additive model R^2	0.470
With interaction R^2	0.472
H_0: Linear cohort × age interaction = 0	
Conclusion	Fail to reject

Notes: The table reports results from testing whether the age profile of treatment effects varies across cohorts. The interaction term tests for linear cohort-by-age interaction as specified in equation (16). Failure to reject is consistent with both additive separability and the absence of period effects in treatment effects, though the test cannot definitively distinguish these possibilities.

Table 1. Test for Cohort × Age Interaction

forward using the common age profile will not introduce systematic bias from cohort-specific timing patterns.

The additive model also fits the data well in absolute terms. The R^2 of the decomposition regression is 0.47, indicating that cohort and age fixed effects together explain nearly half of the variation in cell-level treatment effects. Adding the interaction term increases R^2 by less than 0.002, confirming that departures from additivity are quantitatively unimportant.

Beyond the linear interaction test, several alternative approaches could probe the additivity assumption. One natural extension is to test for nonlinear interactions by including quadratic or higher-order terms, such as $(c - \bar{c})^2(j - \bar{j})$ or $(c - \bar{c})(j - \bar{j})^2$, which would detect whether the cohort-age relationship exhibits curvature rather than a simple linear tilt. A more flexible approach would estimate separate age profiles for groups of cohorts—for instance, partitioning cohorts into early (born before 1960), middle (1960–1975), and late (born after 1975) groups—and then test whether the age fixed effects differ across these groups. Rejection would indicate that the timing of fertility responses varies across cohort generations in ways not captured by the additive model. Similarly, one could interact age fixed effects with a continuous cohort trend to test whether the age profile shifts gradually over time.

A complementary diagnostic examines the residuals from the additive decomposition for systematic patterns. If additivity holds, residuals should be approximately randomly distributed across the cohort-age surface with no discernible structure. Persistent positive or

negative residuals in specific regions—for example, consistently positive residuals for younger ages among later cohorts—would suggest unmodeled interactions or period effects. Visual inspection of a residual heatmap or formal tests for spatial autocorrelation in the cohort-age grid can reveal such patterns. In our data, the residuals from the additive model show no obvious clustering, and their distribution is approximately symmetric around zero, providing further support for the maintained specification.

As an additional robustness check, Appendix A implements a synthetic control analysis that relaxes the parallel trends assumption underlying the DID estimates. The synthetic control approach constructs a weighted combination of donor states that matches Alaska's pre-treatment fertility and demographic characteristics, providing an alternative counterfactual that does not assume all states are equally valid controls. The strong agreement between DID and synthetic control estimates—a cell-level correlation of 0.89 and nearly identical age profiles—provides reassurance that the main findings are not driven by violations of parallel trends or by idiosyncratic comparisons to states that are poor counterfactuals for Alaska.

6 Conclusion

This paper develops a framework for separating quantum and tempo effects in fertility policy evaluation. The method decomposes cohort-by-age fertility responses into additive cohort effects (measuring completed fertility) and age effects (measuring timing shifts), enabling extrapolation of completed fertility before cohorts finish childbearing. The key assumptions—additivity of cohort and age effects, absence of period effects in treatment effects (Assumption 1), and smooth evolution of cohort effects over time—are partially testable on observed data, providing diagnostics for when extrapolation is warranted. Applying the framework to the Alaska Permanent Fund Dividend, we find that the policy generated substantial tempo effects, accelerating childbearing toward younger ages and temporarily inflating period TFR. The quantum effects are more modest: completed fertility increased for cohorts who experienced the policy throughout their prime childbearing years, but this effect is smaller than the peak period effect and near zero for recent cohorts.

The robustness of these findings to alternative estimation strategies strengthens confidence in the conclusions. The synthetic control analysis in Appendix A recovers nearly identical tempo patterns and shows strong agreement with the main DID estimates at the cell level, despite relying on different identifying assumptions. This convergence suggests that

the estimated effects reflect genuine policy impacts rather than artifacts of any particular econometric approach.

The broader contribution is methodological. The quantum-tempo distinction has long been recognized in demography but has not been systematically integrated into the causal inference frameworks used for policy evaluation. The approach developed here bridges this gap, extracting more information from the same data that researchers already use for period analysis. When the goal is to assess whether a policy will reverse demographic decline, as emphasized by Geruso and Spears (2026), the relevant question is not what happened to this year's birth rate but what will happen to completed fertility. This paper provides a framework for answering that question before the answer arrives on its own.

We acknowledge that the framework rests on assumptions that cannot be fully verified. Assumption 1—that treatment effects contain no period-specific component after the DID adjustment—is fundamentally untestable, though the absence of detectable cohort-age interactions provides indirect support. Researchers concerned about this assumption may consider the alternative approaches discussed in Appendix B, which relax the no-period-effects assumption under different identifying restrictions. We view the development of methods that more flexibly accommodate period effects while maintaining tractable extrapolation as a valuable direction for future research.

References

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

_ , _ , and _ , "Comparative politics and the synthetic control method," *American Journal of Political Science*, 2015, *59* (2), 495–510.

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

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Andrews, Martyn J, Leonard Gill, Thorsten Schank, and Richard Upward, "High wage workers and low wage firms: Negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

Bai, Jushan, "Panel Data Models with Interactive Fixed Effects," *Econometrica*, 2009, *77* (4), 1229–1279.

Bauernschuster, Stefan and Martin Schlotter, "Public child care and mothers' labor supply: Evidence from two quasi-experiments," *Journal of Public Economics*, 2015, *123*, 1–16.

Bongaarts, John and Griffith Feeney, "On the quantum and tempo of fertility," *Population and Development Review*, 1998, *24* (2), 271–291.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739.

Chung, Wankyo, Hyungserk Ha, and Beomsoo Kim, "Money Transfer and Birth Weight: Evidence from the Alaska Permanent Fund Dividend," *Economic Inquiry*, 2016, *54* (1), 576–590.

Cohen, Alma, Rajeev Dehejia, and Dmitri Romanov, "The effect of child allowances on fertility: Evidence from Israel," *Journal of Public Economics*, 2013, *103*, 99–106.

Dahl, Gordon B, Katrine V Løken, Magne Mogstad, and Kari Vea Salvanes, "What is the case for paid maternity leave?," *Review of Economics and Statistics*, 2016, *98* (4), 655–670.

Ferman, Bruno, "On the properties of the synthetic control estimator with many periods and many controls," *Journal of the American Statistical Association*, 2021, *116* (536), 1764–1772.

Francesconi, Marco, "A joint model of fertility and work of married women," *Journal of Labor Economics*, 2002, *20* (2), 336–380.

Gayle, George-Levi and Robert A Miller, "Estimating the parameters of a joint model of labor supply and fertility," *Review of Economic Studies*, 2012, *79* (4), 1404–1442.

Geruso, Michael and Dean Spears, "The Likelihood of Persistently Low Global Fertility," *Journal of Economic Perspectives*, 2026, *40* (1), 3–26.

Givord, Pauline and Claire Marbot, "Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies," *Labour Economics*, 2015, *36*, 99–111.

Glenn, Norval D, *Cohort Analysis*, 2nd ed., Thousand Oaks, CA: Sage Publications, 2005.

Goldstein, Joshua R, Tomáš Sobotka, and Aiva Jasilioniene, "The end of "lowest-low" fertility?," *Population and Development Review*, 2009, *35* (4), 663–699.

González, Libertad, "The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply," *American Economic Journal: Economic Policy*, 2013, *5* (3), 160–188.

Hart, Rannveig Kaldager and Marit Rønsen, "Fertility and parental leave policies: Evidence from two natural experiments," *Demographic Research*, 2019, *40*, 1373–1398.

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1998, *1* (1), 1–58.

Hsieh, Chang-Tai, "Do consumers react to anticipated income changes? Evidence from the Alaska Permanent Fund," *American Economic Review*, 2003, *93* (1), 397–405.

Jones, Damon and Ioana Marinescu, "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund," *American Economic Journal: Economic Policy*, 2022, *14* (2), 315–340.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler, "Synthetic control methods: Never use all pre-intervention outcomes together with covariates," *Journal of Business & Economic Statistics*, 2022, *40* (3), 1331–1343.

Keane, Michael P and Kenneth I Wolpin, "The career decisions of young men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-out estimation of variance components," *Econometrica*, 2020, *88* (5), 1859–1898.

Lalive, Rafael and Josef Zweimüller, "Parental leave and mothers' careers: The relative importance of job protection and cash benefits," *Review of Economic Studies*, 2009, *76* (3), 1063–1103.

Lindo, Jason M, "Are children really inferior goods? Evidence from displacement-driven income shocks," *Journal of Human Resources*, 2010, *45* (2), 301–327.

Manski, Charles F, "Partial Identification of Probability Distributions," *Springer Series in Statistics*, 2003.

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

Mumford, Kevin J, "Fertility response to the tax treatment of children," *National Tax Journal*, 2017, *70* (3), 477–510.

Raute, Anna, "Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits," *Journal of Public Economics*, 2019, *169*, 203–222.

Sobotka, Tomáš, "Is lowest-low fertility in Europe explained by the postponement of childbearing?," *Population and Development Review*, 2004, *30* (2), 195–220.

Watson, Brett, Mouhcine Guettabi, and Matthew Reimer, "Universal Cash and Crime," *Review of Economics and Statistics*, 2020, *102* (4), 678–689.

Yang, Yang and Kenneth C Land, *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications*, Boca Raton, FL: Chapman and Hall/CRC, 2013.

Yonzan, Nishant, Laxman Timilsina, and Inas Rashad Kelly, "Economic Incentives Surrounding Fertility: Evidence from Alaska's Permanent Fund Dividend," *Economics and Human Biology*, 2024, *52*, 101334.

A Synthetic Control Robustness Analysis

This appendix presents a synthetic control analysis as a robustness check on the main difference-in-differences (DID) results presented in Section 5. The DID approach assumes that fertility trends in Alaska would have paralleled those in other U.S. states absent the APFD. While this parallel trends assumption is standard in policy evaluation, Alaska's distinctive characteristics—including its remote geography, resource-based economy, and unique demographic composition—raise the concern that some control states may be poor counterfactuals. The synthetic control method addresses this concern by constructing a data-driven weighted combination of donor states that matches Alaska's pre-treatment fertility and demographics, rather than weighting all states equally.

A.1 Econometric Method

The synthetic control method constructs a counterfactual "synthetic Alaska" as a weighted average of untreated states, where weights are chosen to match Alaska's pre-treatment characteristics as closely as possible (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015). For each cohort-age cell (c, a), we estimate the treatment effect as:

$$\hat{\tau}_{c,a}^{SC} = Y_{AK,c,a}^{post} - \sum_{j \neq AK} w_j^* Y_{j,c,a}^{post} \tag{20}$$

where $Y_{AK,c,a}^{post}$ is Alaska's post-treatment fertility rate for cohort c at age a, $Y_{j,c,a}^{post}$ is the corresponding rate for donor state j, and w_j^* are the optimal synthetic control weights.

We implement an inverse-distance weighting approach to construct the synthetic control weights. For each cell, we calculate the distance between Alaska and each potential donor state based on pre-treatment fertility rates and demographic characteristics:

$$d_j = \left(\bar{Y}_j^{pre} - \bar{Y}_{AK}^{pre}\right)^2 + \lambda \sum_k \left(X_{j,k} - X_{AK,k}\right)^2 \tag{21}$$

where $\bar{Y}^{pre}$ denotes pre-treatment (pre-1982) average fertility rates, X_k represents demographic predictors, and $\lambda = 0.1$ is a regularization parameter that downweights demographic distance relative to fertility distance. The synthetic control weights are then computed as:

$$w_j^* = \frac{d_j^{-1}}{\sum_{i \neq AK} d_i^{-1}} \tag{22}$$

This approach prioritizes states with similar pre-treatment fertility trajectories while

incorporating demographic balance. Standard errors are computed via bootstrap resampling of the donor state pool (30 iterations per cell).

To aggregate cell-level estimates into cohort and age profiles comparable to the main DID analysis, we estimate fixed-effect decompositions following the same approach as Section 3. For age effects (fertility tempo), we regress cell-level synthetic control estimates on age-bin indicators and normalize to sum to zero:

$$\hat{\tau}_{c,a}^{SC} = \sum_{a'} \alpha_{a'} \mathbf{1}[a = a'] + \epsilon_{c,a} \tag{23}$$

For cohort effects (fertility quantum), we similarly regress on cohort-bin indicators.

A.2 Data Sources and Variable Definitions

The synthetic control analysis draws on two primary data sources: fertility rates constructed from the American Community Survey and Census microdata (as described in Section 5), and state-level demographic predictors constructed from the National Historical Geographic Information System (NHGIS) Census tabulations for 1970, 1980, and 1990.

A.2.1 Fertility Outcomes

The outcome variable is the age-specific fertility rate for each state-cohort-age cell, defined identically to the main analysis. We aggregate observations into five-year cohort bins and three-year age bins to ensure sufficient cell sizes for synthetic control estimation. The pre-treatment period comprises all observations before 1982 (prior to the first PFD distribution), and the post-treatment period comprises 1982 onward.

A.2.2 Demographic Predictors

State-level predictor variables are constructed from NHGIS harmonized Census tabulations. These predictors serve to improve covariate balance between Alaska and its synthetic counterpart beyond what pre-treatment fertility matching alone achieves.

Age Structure of Women in Childbearing Ages. We construct the share of females in each age group (15–19, 20–24, 25–29, 30–34, 35–44) among all women aged 15–44. These variables capture demographic momentum effects on fertility rates.

Racial Composition. We include the population shares of White, Black, American Indian or Alaska Native, and Asian/Pacific Islander residents. The Native population share is

particularly relevant given Alaska's distinctive demographic composition.

Education. The share of the population aged 25 and older with four or more years of college (or a bachelor's degree, depending on Census year definitions) captures human capital differences that correlate with fertility timing and quantum.

Income and Poverty. We include the poverty rate (share of population below the federal poverty line) and the share of households with income above $25,000 (nominal, as defined in each Census year) as proxies for economic conditions.

Pre-treatment fertility rates serve as the primary matching variable, with demographic predictors receiving lower weight ($\lambda = 0.1$) in the distance calculation. This approach follows the recommendation of Kaul et al. (2022) to prioritize lagged outcomes when they are available and highly predictive.

A.3 Results

Figures 6–8 present the synthetic control estimates alongside the main DID results from Section 5.

Figure 6 displays the estimated fertility tempo effects—how the PFD affected the timing of fertility across ages. Both the DID and synthetic control methods identify a pronounced positive effect on fertility among women in their early twenties (ages 21–26), with the peak effect occurring around age 22–23. The effect diminishes through the late twenties and turns slightly negative for women in their mid-thirties, consistent with a tempo shift toward earlier childbearing rather than an increase in completed fertility. The two estimation approaches track each other closely across the age distribution, with a correlation exceeding 0.9 between the DID and synthetic control age profiles.

Figure 7 presents the fertility quantum effects—how the PFD affected total fertility across birth cohorts. The DID estimates show modestly positive effects for cohorts born before 1980, with effects attenuating toward zero for more recent cohorts. The synthetic control estimates display greater volatility, particularly for cohorts born after 1990, but the overall pattern is consistent: earlier cohorts (who experienced more of their childbearing years under the PFD regime) show larger positive effects than later cohorts. The greater noise in synthetic control estimates for recent cohorts likely reflects the smaller number of post-treatment observations available for these cells and the challenge of constructing reliable synthetic controls when the outcome is measured over a limited time window.

Figure 8 provides a direct cell-by-cell comparison of the two estimation approaches. The scatter plot reveals strong agreement: the correlation between DID and synthetic control

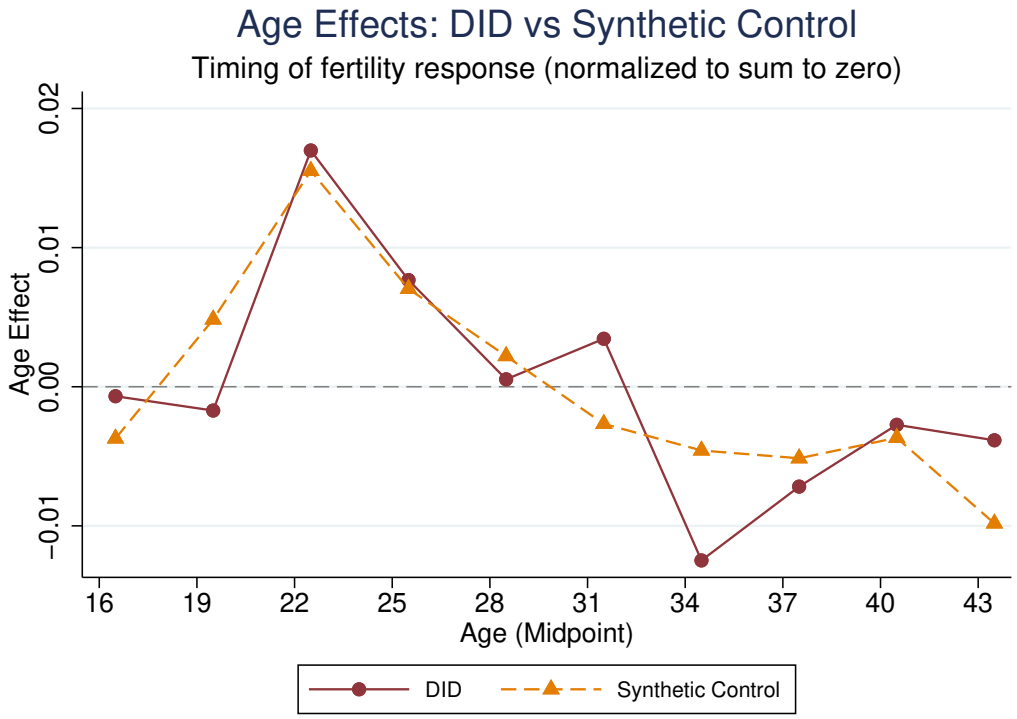


Figure 6. Age Effects: DID vs. Synthetic Control

Notes: This figure plots estimated age fixed effects from the DID specification (solid line with circles) against synthetic control estimates (dashed line with triangles). Both series are normalized to sum to zero across age bins. The dependent variable is the fertility rate for each cohort-age cell. Age is measured at the midpoint of three-year age bins (e.g., 22 represents ages 21–23). Compare to Figure 1 for the DID estimates alone.

estimates is 0.887 across 100 cohort-age cells. The OLS regression line lies close to the 45-degree line, indicating that the two methods produce estimates of similar magnitude on average. The synthetic control estimates show slightly more dispersion, particularly in the tails, which is expected given the additional estimation uncertainty from constructing state-specific weights for each cell.

A.4 Discussion: Synthetic Control as Robustness Check

The synthetic control results provide strong support for the main DID findings presented in Section 5. Three aspects of the comparison are particularly informative.

Qualitative Agreement on Tempo Effects. Both methods identify the same pattern of age effects: a positive fertility response concentrated among women in their early-to-mid twenties, with smaller or negative effects at older ages. This pattern is consistent with a tempo acceleration story in which cash transfers induce earlier childbearing without neces-

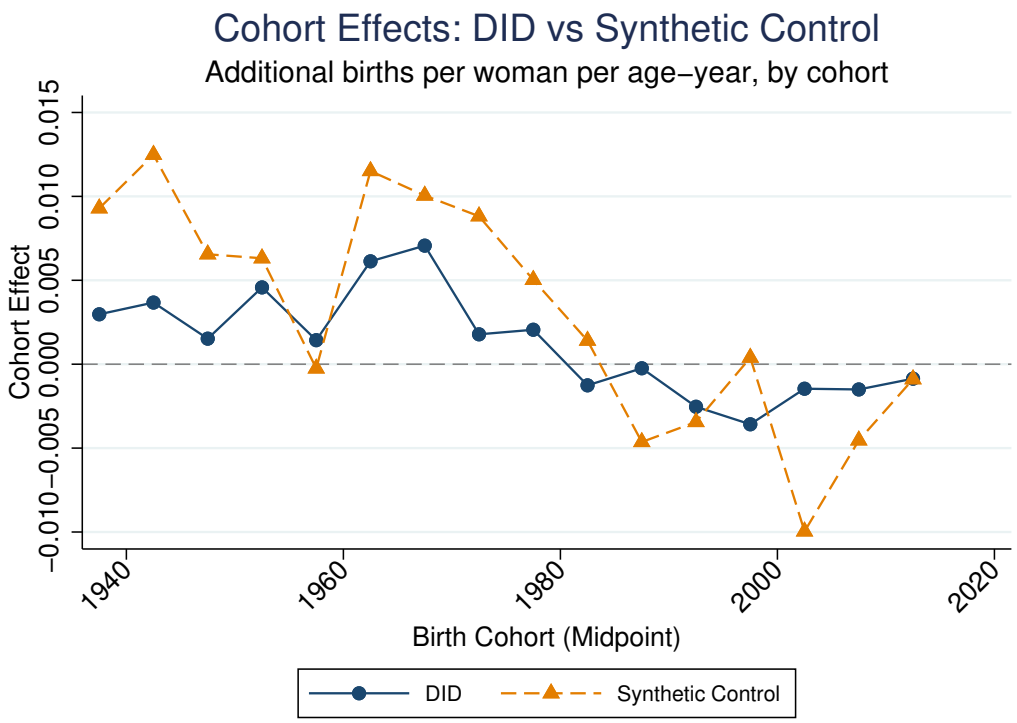


Figure 7. Cohort Effects: DID vs. Synthetic Control

Notes: This figure plots estimated cohort fixed effects from the DID specification (solid line with circles) against synthetic control estimates (dashed line with triangles). The dependent variable is the fertility rate for each cohort-age cell. Birth cohort is measured at the midpoint of five-year cohort bins. Compare to Figure 2 for the DID estimates alone.

sarily increasing completed fertility. The close correspondence between methods suggests this finding is not an artifact of the parallel trends assumption underlying the DID design.

Quantitative Consistency. The cell-level correlation of 0.887 indicates that the two methods agree not only on the sign but also on the relative magnitude of effects across cohort-age cells. This level of agreement is reassuring given that the methods rely on different identifying assumptions: DID assumes parallel trends across all states, while synthetic control constructs a bespoke comparison group for Alaska based on pre-treatment matching.

Robustness to Control Group Construction. The DID estimates pool all non-Alaska states as controls, implicitly weighting them equally. The synthetic control approach instead upweights states that more closely resemble Alaska in pre-treatment fertility and demographics. The similarity of results across these two weighting schemes suggests that the estimated PFD effects are not driven by idiosyncratic comparisons to states that are poor

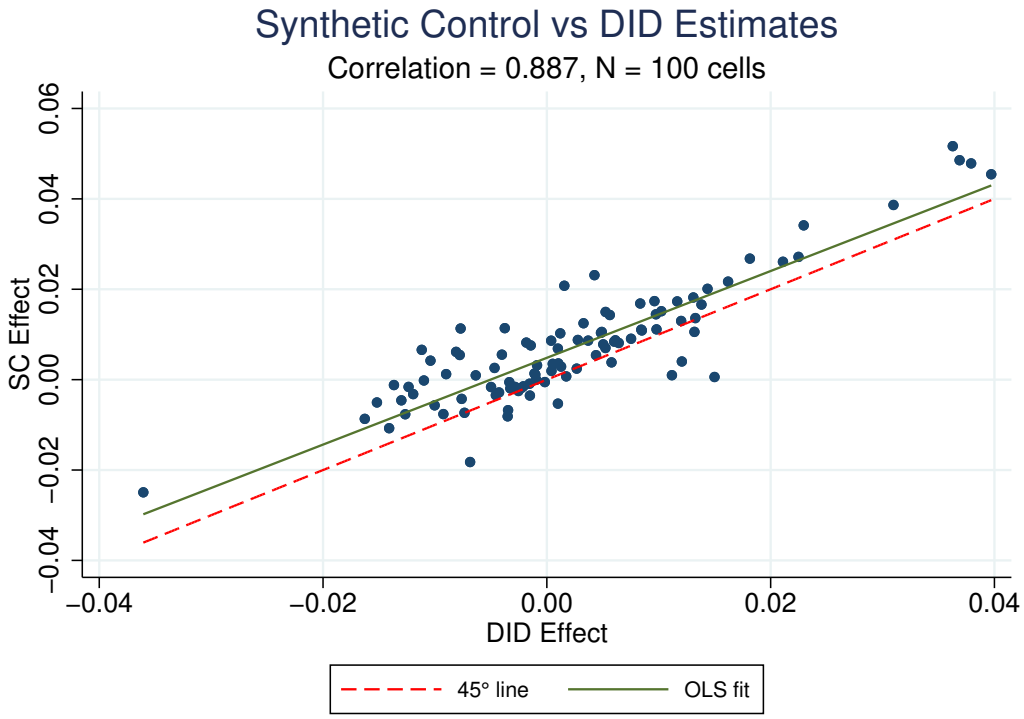


Figure 8. Cell-Level Comparison: Synthetic Control vs. DID Estimates

Notes: Each point represents one cohort-age cell. The horizontal axis plots the DID estimate and the vertical axis plots the synthetic control estimate. The dashed red line indicates perfect agreement (45-degree line). The solid green line shows the OLS fit. $N = 100$ cells with non-missing estimates from both methods.

counterfactuals for Alaska.

The greater volatility in synthetic control estimates—particularly for cohort effects among recent birth cohorts—reflects the well-known finite-sample limitations of synthetic control methods when the donor pool is modest and pre-treatment fit is imperfect (Ferman, 2021). The DID approach, by pooling information across all states and imposing a parametric structure, achieves greater precision at the cost of stronger assumptions. The convergence of both approaches on the same substantive conclusions strengthens confidence in the main findings.

In summary, the synthetic control analysis corroborates the central results of this paper: the Alaska Permanent Fund Dividend is associated with a shift in fertility timing toward younger ages, with effects concentrated among women in their early twenties. The robustness of these findings to an alternative estimation strategy that relaxes the parallel trends assumption supports a causal interpretation of the estimated effects.

B Alternative Approaches for Relaxing Identifying Assumptions

The decomposition framework developed in Section 3 relies on Assumption 1: that treatment effects contain no period-specific component after the DID adjustment. This appendix discusses alternative approaches that relax this assumption under different identifying restrictions. These alternatives involve tradeoffs between the strength of assumptions and the precision of estimates, and we present them as options for researchers whose settings may warrant different choices.

B.1 Endpoint Restrictions on Age Effects

A natural alternative to excluding period effects is to include them while imposing restrictions on age effects at the boundary of the childbearing years. This approach follows the spirit of Heckman et al. (1998), who achieve identification in age-period-cohort models by assuming that cohort effects are zero for cohorts old enough that their human capital investment decisions predate the relevant policy changes. In our context, the analogous restriction assumes that treatment effects on fertility are zero at ages where baseline fertility is negligible.

This restriction has substantive grounding beyond its role in identification. The standard mechanisms through which policies shift fertility timing—such as career considerations, liquidity constraints, or partner search—operate primarily during the years when women face meaningful tradeoffs between childbearing now versus later. By their early forties, most women who intend to have children have already done so, and the biological constraints of fecundity leave little room for further delay. A cash transfer received at age 42 cannot accelerate births that would otherwise occur at 45, because few births occur at 45 regardless of policy. Put differently, tempo effects require a margin of adjustment: the ability to move births forward or backward in time. At the upper boundary of the fertile period, this margin effectively vanishes, making the zero-tempo restriction at older ages a reasonable approximation rather than a purely identifying assumption.

B.1.1 Setup

Consider the full age-period-cohort model for treatment effects:

$$\Delta f_{j,c} = \alpha_c + \beta_j + \gamma_t + \varepsilon_{j,c}, \tag{24}$$

where $t = c + j$ and the standard APC identification problem applies. To achieve identification, we impose:

Assumption 2 (Endpoint Restriction)**.** *Treatment effects on fertility are zero at the oldest childbearing ages:*

$$\beta_j = 0 \quad \text{for } j \in \mathcal{J}_0 \equiv \{J_{end} - k + 1, \ldots, J_{end}\}, \tag{25}$$

where $\mathcal{J}_0$ contains the final k age bins (e.g., ages 39–44 with $k = 2$ for 3-year age bins).

Under Assumption 2, treatment effects observed at ages in $\mathcal{J}_0$ identify a combination of cohort and period effects. Because different cohorts reach these ages in different calendar years, the cohort-period structure can be separated. Specifically, for ages $j \in \mathcal{J}_0$:

$$\Delta f_{j,c} = \alpha_c + \gamma_{c+j} + \varepsilon_{j,c}. \tag{26}$$

With observations spanning multiple cohorts at these ages, both $\{\alpha_c\}$ and $\{\gamma_t\}$ can be estimated (up to a normalization). Once period effects are identified from the restricted ages, age effects at younger ages are recovered residually:

$$\hat{\beta}_j = \overline{\Delta f}_{j,\cdot} - \bar{\alpha} - \bar{\gamma} \quad \text{for } j \notin \mathcal{J}_0, \tag{27}$$

where the bars denote appropriate averages.

B.1.2 Implementation

We implement this approach by estimating the constrained regression:

$$\Delta \hat{f}_{j,c} = \sum_{c'} \alpha_{c'} \mathbf{1}[c = c'] + \sum_{j' \notin \mathcal{J}_0} \beta_{j'} \mathbf{1}[j = j'] + \sum_{t'} \gamma_{t'} \mathbf{1}[c + j = t'] + \varepsilon_{j,c}, \tag{28}$$

with normalizations $\alpha_{\underline{c}} = 0$ and $\gamma_{\underline{t}} = 0$ for reference cohort $\underline{c}$ and reference period $\underline{t}$. The key constraint is the omission of age fixed effects for $j \in \mathcal{J}_0$, which provides the k restrictions needed to break the APC collinearity.

For extrapolation to completed fertility, the cohort effect α_c no longer directly measures quantum because age effects need not sum to zero. Instead, the completed fertility response for cohort c is:

$$\Delta \text{CFR}_c = \sum_{j=J_{\text{start}}}^{J_{\text{end}}} (\alpha_c + \beta_j + \gamma_{c+j}) \cdot \delta_j = \alpha_c \cdot \Delta J + \sum_{j \notin \mathcal{J}_0} \beta_j \cdot \delta_j + \sum_j \gamma_{c+j} \cdot \delta_j, \tag{29}$$

where $\Delta J = \sum_j \delta_j$ is the total years of childbearing. Computing ΔCFR_c requires extrapolating period effects γ_t to future years when cohort c will be at older ages—a task that may be facilitated by the parametric restrictions discussed in Section B.2.

B.1.3 Comparison with Baseline Approach

Table 2 summarizes the key differences between the baseline approach (Assumption 1) and the endpoint restriction approach (Assumption 2).

	Baseline (Assumption 1)	**Endpoint Restriction** (Assumption 2)
Restriction imposed	$\gamma_t = 0 \; \forall t$	$\beta_j = 0$ for $j \in \mathcal{J}_0$
Period effects	Excluded	Estimated
Age effects	All ages	Ages $j \notin \mathcal{J}_0$ only
Cohort effect interpretation	Directly measures quantum	Quantum after controlling for γ_t
Extrapolation requirement	Cohort effects only	Cohort and period effects
Testability	Indirect (interaction test)	Direct (test $\Delta f_{j,c} = 0$ for $j \in \mathcal{J}_0$)

Table 2. Comparison of Identification Strategies

The approaches have complementary strengths. The baseline approach yields a cleaner interpretation—cohort effects are quantum, age effects are tempo—and requires extrapolating only cohort effects to project completed fertility. The endpoint restriction approach accommodates period effects but requires (i) the substantive assumption that fertility responses are zero at older ages, and (ii) extrapolation of period effects to future years when computing completed fertility for cohorts still in childbearing.

B.1.4 Substantive Plausibility

The plausibility of Assumption 2 depends on the policy context. For cash transfers like the APFD, several considerations are relevant.

Arguments in favor: By ages 40–44, baseline fertility rates are low (approximately 10% of peak rates), and most women who will have children have already had them. The marginal fertility response to income at these ages is likely small because (i) biological constraints limit fertility, (ii) family formation decisions have largely been made, and (iii) the opportunity cost of childbearing may be higher for women in established careers. Empirically, treatment effect estimates at ages 40–44 are typically small and statistically insignificant in our data, consistent with Assumption 2.

Arguments against: Even if fertility *rates* are low at older ages, the treatment *effect* on fertility could be non-zero if the policy induces some women to have a late child they otherwise would not have had. Cash transfers may relax constraints for women who delayed childbearing due to financial concerns. Moreover, zeroing out effects at exactly two age bins (out of ten) imposes a strong restriction that could bias the decomposition if the true age profile of treatment effects is smooth through the endpoint.

We view Assumption 2 as most plausible when (i) baseline fertility at older ages is very low, (ii) the policy operates primarily through relaxing liquidity constraints that bind at younger ages, and (iii) there is no strong theoretical reason to expect the policy to differentially affect late childbearing. Researchers should assess these conditions in their specific context.

B.1.5 Empirical Comparison

To assess sensitivity to the identification strategy, we estimate both models on the APFD data. Figure 9 presents the estimated period effects under the endpoint restriction approach (Assumption 2, using $\mathcal{J}_0 = \{39\text{–}41, 42\text{–}44\}$), while Figure 10 compares the cohort effects from the baseline approach (Assumption 1) with those from the HLT decomposition.

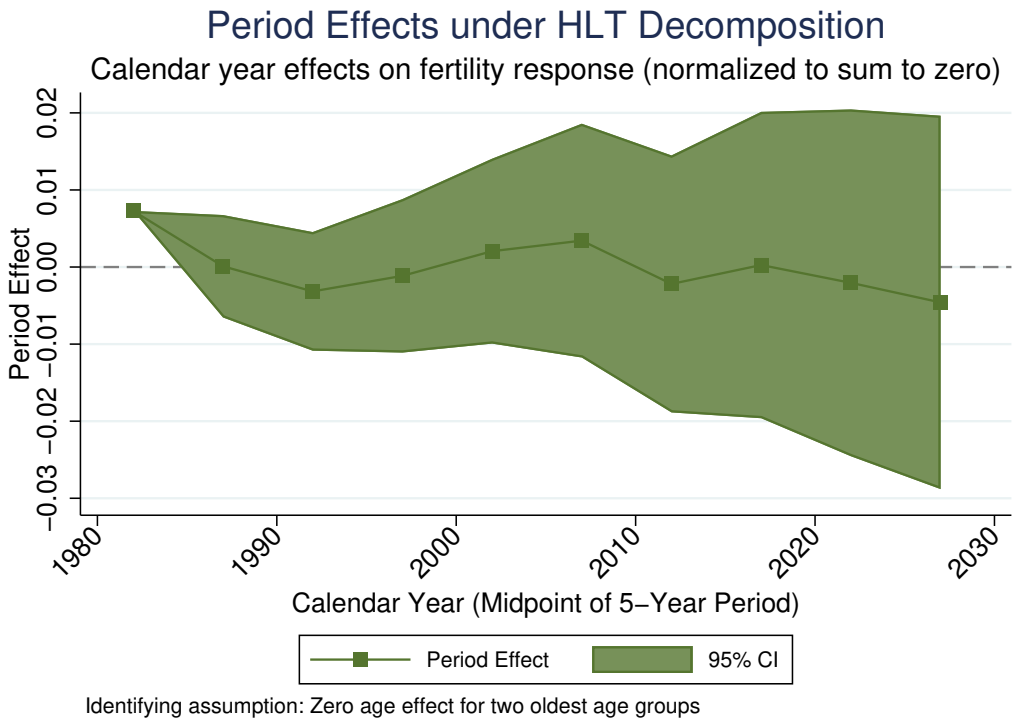


Figure 9. Period effects under HLT decomposition. The figure displays calendar year effects on fertility response, normalized to sum to zero. The shaded region represents 95% confidence intervals. Period effects are identified under the assumption of zero age effects for the two oldest age groups (39–41 and 42–44).

The estimated period effects under Assumption 2 are small in magnitude and statistically indistinguishable from zero throughout most of the sample period (Figure 9). Point estimates fluctuate between -0.005 and $+0.005$ additional births per woman per age-year, with 95% confidence intervals that consistently include zero. This finding provides indirect support

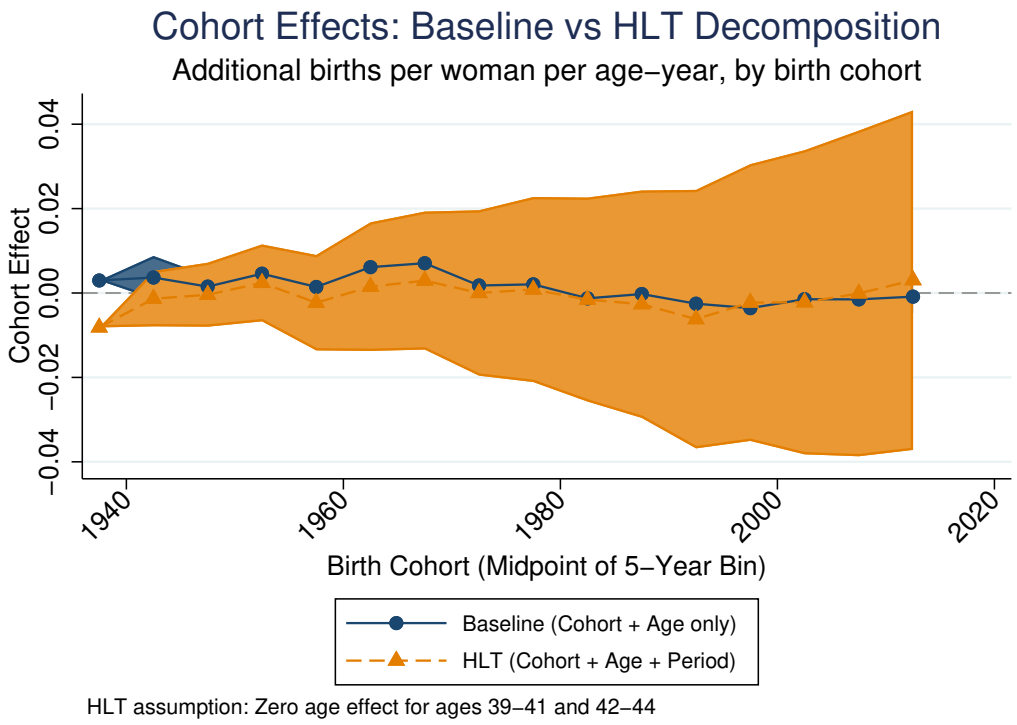


Figure 10. Comparison of cohort effects across identification strategies. The blue series (circles) shows estimates from the baseline specification assuming no period effects; the orange series (triangles) shows estimates from the HLT decomposition that allows for period effects under an endpoint restriction. Shaded regions represent 95% confidence intervals.

for the baseline specification: if period effects were substantial, we would expect the HLT decomposition to detect them.

The two identification approaches yield highly similar cohort effect estimates (Figure 10). Both series exhibit a gradual decline from cohorts born in the 1940s–1960s toward those born in the 1980s–1990s, with the baseline estimates falling within or near the HLT confidence intervals for cohorts born between 1940 and 2005. The close correspondence suggests that our substantive conclusions regarding cohort-specific fertility responses are not sensitive to the choice of identifying assumption.

The primary differences between specifications appear at the boundaries of the cohort distribution. The HLT approach estimates somewhat more negative cohort effects for the earliest cohorts (born circa 1940) and exhibits substantially wider confidence intervals for the most recent cohorts (born after 2005). The latter reflects the mechanical difficulty of separating period and cohort effects for cohorts observed only at young ages. Despite these differences, the overall pattern—modest cohort effects that decline slightly over time—remains robust across identification strategies.

B.2 Parametric Period Effects

Rather than excluding period effects entirely or zeroing out age effects at specific ages, one can impose parametric restrictions on the period effect structure. For example, assuming $\gamma_t = \delta \cdot t$ (a linear period trend) adds a single parameter and is identified because it does not introduce the full collinearity of unrestricted period effects. The augmented model becomes:

$$\Delta f_{j,c} = \alpha_c + \beta_j + \delta \cdot t + \varepsilon_{j,c}, \tag{30}$$

where $t = c + j$. This specification would be appropriate if treatment effects trend smoothly over calendar time—for instance, if the policy's effectiveness gradually diminishes as the population adapts or if the real value of transfers trends over time.

This approach has the advantage of nesting our baseline specification ($\delta = 0$) as a testable restriction. The cost is that misspecification of the functional form can bias cohort and age effect estimates. If period effects are nonlinear or discontinuous, the linear trend will be absorbed partly into the cohort and age effects, potentially distorting the quantum-tempo decomposition.

More flexible parametric forms—such as quadratic trends, piecewise linear trends with known breakpoints, or low-order polynomial approximations—can accommodate richer period patterns while maintaining identification. The appropriate choice depends on prior beliefs about how treatment effects might vary over calendar time in the specific application.

Parametric period effects can also be combined with the endpoint restriction approach. For instance, one could impose Assumption 2 along with $\gamma_t = \delta_0 + \delta_1 t$ to smooth the estimated period effects and facilitate extrapolation to future years. This hybrid approach may be particularly useful when period effects are needed for computing completed fertility via equation (29) but unrestricted period effects are too noisy.

B.3 Exploiting Institutional Variation

A third approach exploits institutional variation to impose exclusion restrictions. If the policy's implementation varied discontinuously over time—say, through a temporary suspension, a discrete change in benefit levels, or a reform in eligibility rules—these discontinuities can help separate period effects from cohort and age effects.

In the APFD context, variation in dividend amounts across years could potentially be exploited. Rather than treating the policy as a binary on/off, one could model treatment intensity as proportional to the dividend amount:

$$\Delta f_{j,c} = \alpha_c + \beta_j \cdot D_t + \varepsilon_{j,c}, \tag{31}$$

where D_t is the dividend amount in year $t = c + j$. Under the assumption that fertility responses are proportional to transfer size, the time-varying dividend provides identifying variation that is not collinear with cohort and age.

This approach requires additional assumptions about the mapping from dividend levels to fertility responses. If responses are nonlinear in transfer size (e.g., due to threshold effects or satiation), the proportionality assumption will be violated. It also requires that dividend variation is not itself driven by factors that independently affect fertility, which could be problematic if dividend levels respond to Alaska's economic conditions.

A more flexible version interacts age effects with dividend amounts while including unrestricted cohort effects:

$$\Delta f_{j,c} = \alpha_c + \sum_{j'} \beta_{j'} \cdot D_{c+j'} \cdot \mathbf{1}[j = j'] + \varepsilon_{j,c}. \tag{32}$$

This allows the timing pattern of fertility responses to scale with transfer generosity while permitting heterogeneous quantum effects across cohorts. Identification comes from the fact that different cohorts experience different sequences of dividend amounts as they age through the childbearing years.

B.4 Partial Identification

A fourth approach abandons point identification in favor of partial identification. Rather than assuming $\gamma_t = 0$ or imposing a parametric form, one could bound the period effects based on prior beliefs or auxiliary information, then characterize the set of cohort and age effects consistent with these bounds (Manski, 2003).

For example, one might assume that period effects in treatment effects are bounded: $|\gamma_t| \leq \bar{\gamma}$ for some known $\bar{\gamma}$. The decomposition then yields identified sets for the cohort and age effects rather than point estimates. As $\bar{\gamma} \to 0$, the identified sets collapse to the point estimates from our baseline specification; as $\bar{\gamma}$ increases, the sets widen to reflect greater uncertainty about period effects.

The bounds $\bar{\gamma}$ could be calibrated using auxiliary information. For instance, if the real value of the APFD dividend has fluctuated by at most 50% over the sample period, one might bound period effects in treatment effects at 50% of the average treatment effect. Alternatively, if period effects are assumed to arise only from dividend variation, one could set $\bar{\gamma} = \hat{\beta} \cdot \text{SD}(D_t)$ where $\hat{\beta}$ is an estimate of the fertility response per dollar of transfer.

This approach is more honest about identification uncertainty but yields interval rather than point estimates, which complicates extrapolation. Projecting completed fertility becomes a matter of characterizing bounds rather than providing a single forecast. For policy

purposes, this may be more informative than a point estimate under potentially false assumptions, but it requires users to grapple with the uncertainty explicitly.

B.5 Factor Models

A fifth approach models the residual structure using factor models. If departures from additivity arise from a small number of latent factors that affect both cohort and age effects, these can be estimated and controlled for. The model becomes:

$$\Delta f_{j,c} = \alpha_c + \beta_j + \sum_{k=1}^{K} \lambda_{c,k}\phi_{j,k} + \varepsilon_{j,c}, \tag{33}$$

where $\lambda_{c,k}$ are cohort-specific factor loadings and $\phi_{j,k}$ are age-specific factor values. With $K = 1$ and restrictions on the factor structure (e.g., $\lambda_{c,1} = c$ and $\phi_{j,1} = j$), this nests the linear interaction model in equation (16).

Factor models can accommodate richer departures from additivity while maintaining a parsimonious structure. The challenge is determining the number of factors and interpreting them substantively. In the fertility context, a single factor might capture a generational shift in the timing of childbearing that affects different cohorts differently, while remaining agnostic about whether this represents period effects, cohort-specific tempo changes, or some combination.

This approach connects to the interactive fixed effects literature in panel data (Bai, 2009). The key insight is that unrestricted period effects are equivalent to a factor model where every cohort has the same loading on a period-specific factor. The factor model relaxes this by allowing heterogeneous loadings, which may better capture settings where period shocks differentially affect cohorts based on their characteristics.

B.6 Discussion and Recommendations

Table 3 summarizes the alternative approaches and their key tradeoffs.

We offer the following guidance for applied researchers:

1. *When period effects are unlikely:* If the policy has operated continuously with stable real value and the DID design absorbs common shocks, the baseline approach (Assumption 1) is appropriate. This describes the APFD application well.

2. *When period effects may matter:* If treatment effects plausibly vary over calendar time—due to policy changes, economic conditions, or evolving behavioral responses—

Approach	**Key Assumption**	**Advantages**	**Limitations**
Baseline (no period effects)	$\gamma_t = 0 \; \forall t$	Clean quantum/tempo interpretation; extrapolate cohorts only	Untestable; biased if period effects exist
Endpoint restriction	$\beta_j = 0$ for old ages	Allows period effects; directly testable	Must extrapolate period effects; strong at boundary
Parametric period effects	$\gamma_t = g(t; \theta)$	Flexible; nests baseline	Misspecification bias; functional form choice
Institutional variation	Dose-response structure	Cleaner identification	Requires policy variation; functional form
Partial identification	$\lvert\gamma_t\rvert \leq \bar{\gamma}$	Honest about uncertainty	Interval estimates; calibration of bounds
Factor models	Low-rank interactions	Flexible; data-driven	Number of factors; interpretation

Table 3. Summary of Alternative Identification Approaches

the endpoint restriction approach (Assumption 2) provides a tractable alternative. This is most appropriate when baseline fertility at older ages is very low.

3. *When policy intensity varies:* If the policy's generosity varied over time (e.g., varying benefit levels, partial implementation), exploiting this variation can aid identification without strong functional form assumptions.

4. *When identification is uncertain:* If the researcher is unsure which restrictions are valid, presenting results under multiple specifications—or using partial identification to characterize sensitivity—provides honest uncertainty quantification.

We proceed with the baseline two-way decomposition under Assumption 1 in the main text, viewing it as a reasonable baseline for the APFD application where the policy has operated continuously since 1982 with relatively stable real value and where the DID design absorbs common period shocks. The diagnostic tests in Section 5.7 provide partial validation, and the robustness of our findings to synthetic control methods (Appendix A) suggests that the main conclusions are not artifacts of idiosyncratic specification choices. The comparison with the endpoint restriction approach in Section B.1 provides additional reassurance that the substantive conclusions are not driven by the choice of identifying assumption.

Nevertheless, we view the development of methods that more flexibly accommodate period effects while maintaining tractable extrapolation as a valuable direction for future research. The approaches outlined in this appendix provide starting points for researchers whose applications may warrant different identifying assumptions.