

**NATIONAL PENSIONS REGULATORY AUTHORITY**

**When Pension Reform Does Not Bite: Reform Measurement, Informality, and  
Labour Market Responses Across Countries**

Kennedy K. Abrokwa, Victor Azuma Mejida, and Philemon Laar

**WP/26/003**

NPRA Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in NPRA Working Papers are those of the author(s) and do not necessarily represent the views of the National Pensions Regulatory Authority, its Board, or its management.

MAY 2026

## **NPRA Working Paper**

Planning, Research, Monitoring and Evaluation Directorate

Authorized for distribution by Director, PRME

### **Abstract**

This paper examines the macro-labor market impacts of public pension adjustments using a comprehensive global panel of 183 countries from 2000 through 2024. Its central contribution is to demonstrate that estimated policy impacts depend fundamentally on how a reform event is taxonomically defined and on the institutional formalization of the labor market in which it is executed. Utilizing the interaction-weighted event-study design formalized by Sun and Abraham (2021) to insulate estimates against staggered treatment biases, the analysis directly contrasts two alternative reform measures: a broad definition treating any recorded administrative or legislative update as a uniform treatment onset, and a restricted definition that isolates exclusively major structural and parametric shocks that directly alter individual budget constraints. The interactive fixed-effects specifications provide three evidence. First, while both definitions detect a downward trajectory in labor force participation (LFP) within non-African economies, the restricted taxonomy removes institutional noise to yield larger, more economically substantive, and more precisely estimated point estimates, capturing an absolute 4.09 percentage point contraction within four years of enactment compared to an attenuated 3.81 percentage points under the broad classification. Second, across African economies, the estimated treatment effects under both definitions are small, short-lived, and statistically indistinguishable from zero, confirming that low statutory coverage and pervasive labor market informality act as a structural brake that completely decouples behavioral responses from formal policy adjustments. Third, the macro-labor response is highly unequal across alternative outcomes and gender lines. Within non-African settings, declining participation is mirrored by a sharp expansion in the national unemployment rate, verifying a severe demand-side worker displacement friction, while female labor supply drives the overall downward contraction across all operational specifications. Finally, forward-looking counterfactual micro-simulations and never-treated regional projections illustrate the broad economic scale of these parameters, establishing that the measurement of a reform is not a neutral technical choice but a substantive one that alters empirical conclusions. Our results prove that explicitly accounting for taxonomical boundaries and underlying labor market segmentation is a prerequisite for the rigorous econometric evaluation of social protection systems.

**JEL Classification:** H55, J21, J26

**Keywords:** Pension Reform, Labor Force Participation, Staggered Event Study, Heterogeneous Treatment Effects, Informality, Gender Decomposition.

**Author's E-Mail:** [research@npra.gov.gh]

## 1. Introduction

Population ageing is placing increasing pressure on pension systems worldwide, prompting governments to implement a wide range of reforms aimed at improving long-term sustainability. A large empirical literature, primarily based on advanced economies, finds that pension reforms, particularly those affecting retirement age and contribution incentives, influence labour supply decisions (Gruber and Wise, 1999; Staubli and Zweimüller, 2013; Cribb, Emmerson, and Tetlow, 2014). The underlying premise is straightforward: changes in pension rules alter incentives, and individuals adjust their labour market behaviour accordingly.

However, this premise rests on institutional conditions that are not universally present. It implicitly assumes that pension systems have broad coverage, that labour markets are predominantly formal, and that statutory rules are effectively enforced. In many developing and emerging economies, particularly in Africa, these conditions are weak or incomplete. Pension coverage in sub-Saharan Africa remains below 10 percent of the labour force, while informal employment accounts for more than 70 percent of total employment in most countries in the region (ILO, 2017, 2018). Under such conditions, the transmission from pension policy to individual labour supply decisions may be substantially attenuated.

This paper revisits the relationship between pension reform and labour supply by focusing on an overlooked but fundamental issue: whether the way pension reforms are identified and classified in cross-country datasets affects the estimated effects themselves. Widely used datasets, such as the ISSA country profiles, document a broad spectrum of policy activity, spanning foundational legislation, parametric adjustments, administrative amendments, and implementation phases. In most empirical applications, these heterogeneous events are treated equivalently as reforms, despite capturing distinct underlying processes (Holzmann and Hinz, 2005). Treating distinct forms of policy activity as equivalent conflates gradual institutional evolution with discrete policy shocks.

This distinction is consequential. Using a global panel of 183 countries over the period 2000 to 2024 and a staggered difference in differences event study design, two alternative measures of pension reform are constructed. The broad definition treats any recorded reform activity as a treatment event. The restricted definition isolates major structural and parametric reforms that directly alter incentives. The baseline specification follows Sun and Abraham (2021), whose interaction-weighted estimator addresses known biases in two-way fixed effects models under heterogeneous treatment timing (Goodman-Bacon, 2021; De Chaisemartin and D'Haultfoeuille, 2020).

Both definitions produce negative effects on labour force participation in non-African countries, but the restricted definition yields larger and more precisely estimated effects, with a cumulative decline of 4.1 percentage points within four years compared to 3.8 under the broad definition. In African countries, the effects are small, short lived, and statistically indistinguishable from zero under both definitions. These findings

indicate that the same underlying reform events, classified differently, generate different conclusions about the magnitude and precision of labour supply effects, and that institutional context mediates the response.

The paper further decomposes the labour market response along three margins. First, the analysis distinguishes between demand and supply channels by examining LFP and unemployment together. In non-African countries, rising unemployment accompanies the decline in participation, consistent with demand driven displacement. Second, gender decomposition reveals that female labour supply declines more than total participation in every sample and under every definition, indicating that women bear a disproportionate share of the adjustment. Third, counterfactual simulations for treated countries and projections for countries that never reformed during the sample period illustrate the economic magnitude of the estimated effects and connect the empirical findings to practical policy questions.

This paper contributes to the literature in three ways. First, it demonstrates that the estimated labour market effects of pension reform depend critically on how reform is measured, with broader definitions attenuating the estimated response relative to definitions that isolate major structural and parametric changes. Second, it shows that the transmission of pension reform into labour supply is mediated by labour market structure, with significant effects concentrated in countries with formalized labour markets and negligible effects in high-informality settings. Third, it decomposes the labour market response by gender and by the demand versus supply margin, revealing distributional consequences that are masked in aggregate analyses.

The remainder of the paper proceeds as follows. Section 2 reviews the literature. Section 3 describes the data and the construction of alternative reform measures. Section 4 outlines the empirical strategy. Sections 5 and 6 present the results using the broad and restricted reform definitions. Section 7 presents simulations and counterfactual exercises. Section 8 concludes with policy implications.

## **2. Literature Review**

The relationship between pension policy and labour supply is one of the most extensively studied questions in applied labour economics. The theoretical foundation is well established: pension systems influence labour market behaviour through the statutory retirement age, the generosity of benefits, and the effective tax rate on continued work embedded in contribution requirements and benefit formulas. The core prediction, formalised in the option-value framework of Stock and Wise (1990) and the landmark cross-country studies of Gruber and Wise (1999, 2004), is that changes in these parameters alter the implicit tax on work and thereby shift labour supply decisions. When pension rules change, individuals adjust their behaviour, or so the theory predicts.

The empirical evidence from advanced economies has largely confirmed this prediction, though with important nuances. Gruber and Wise (1999), in an influential eleven-country study, demonstrated that the implicit tax on work embedded in pension systems strongly predicts early retirement behaviour across countries. Subsequent work has refined this finding in a range of institutional settings. Staubli and Zweimüller (2013) found that raising the early retirement age in Austria increased employment by approximately 10 percentage points among affected cohorts. In the United Kingdom, Cribb et al. (2013, 2014) examined the increase in the female state pension age from 60 to 62 between 2010 and 2014, finding that employment rates among affected women increased by 7.3 percentage points at age 60, with effects persisting as the pension age rose further. Importantly, they also documented a 4.2 percentage point increase in employment among affected women's male partners, providing evidence of complementarities in leisure within couples. The authors argued that the magnitude and pattern of responses were more consistent with the state pension age acting as a shock or signal, anchoring social norms about the appropriate age to retire, rather than operating purely through financial incentives.

Evidence from Germany further demonstrates that pension reforms affect behaviour well before the point of retirement. Artmann et al. (2023) exploited the 2014 *Mütterrente* reform, which increased pension wealth for mothers of children born before January 1, 1992, by an average of 4.4 percent per child. Using administrative data on the universe of working histories, they documented significant reductions in labour earnings among affected women when they were on average aged 50 to 55, far from retirement. Their estimates imply that an extra euro of pension wealth reduces unconditional labour earnings by 54 cents, a marginal propensity to earn consistent with estimates from lottery winnings. This finding is particularly significant because it demonstrates forward-looking behaviour: individuals adjust labour supply in middle age in response to changes in future pension entitlements.

The Norwegian experience offers complementary insights into the relative effectiveness of alternative reform strategies. Hernæs et al. (2015) exploited the comprehensive 2011 Norwegian pension reform, which removed a confiscatory earnings test for private sector workers while leaving the earliest access age unchanged. They found that the removal of the earnings test increased annual earnings at age 63 by approximately NOK 14,700, roughly 25 percent of pre-reform earnings, implying an uncompensated labour earnings elasticity of around 0.25. Critically, the pure flexibility reform, improving work incentives without changing the access age, produced labour supply responses of similar magnitude to reforms that also increased the access age, while the latter generated additional spillovers into disability insurance claims. This suggests that improving the returns to work can be as effective as mandating longer working lives, with fewer adverse side effects.

Yet the transmission of pension policy into behaviour is far from automatic, even in the advanced economy settings where the evidence is strongest. Propper et al. (2023) examined the response of senior doctors in the English National Health Service to a

major public sector pension reform that moved employees from a final salary to a career average plan. Despite a reduction in pension wealth of approximately 40 percent, doctors increased their labour supply by only about 3.5 percent, implying a labour supply elasticity with respect to pension wealth of just  $-\$0.05$ . The authors attribute this modest response partly to the complexity of the pension scheme and the difficulty individuals face in translating changes in deferred remuneration into current labour supply decisions. This finding underscores a broader point: the behavioural response to pension reform depends not only on the design of the reform but on its salience, the transparency of the policy change, and the extent to which individuals understand and act upon changes in pension incentives.

The evidence base described above is concentrated almost exclusively in advanced economies. This is not coincidental. The identification strategies employed in these studies require institutional conditions that are prevalent in high-income settings: comprehensive administrative data, well-documented reform histories, stable legal frameworks, and, most importantly, pension systems with broad population coverage and labour markets that are predominantly formal. Where these conditions are absent, both the transmission of pension policy into behaviour and the ability to detect such effects empirically are compromised.

The literature on pension reform in developing and emerging economies is considerably thinner and points in divergent directions. Aguila (2011) documented modest labour supply responses to Mexico's pension reform, concentrated among formal sector workers, precisely the group that the reform could reach. Where effects are detected, they tend to be smaller in magnitude and less robust than those found in advanced economies. The divergence is not merely a matter of data quality or methodological rigour. It reflects fundamental structural differences in labour markets that mediate the transmission of pension policy into individual behaviour.

The central mediating factor is informality. In economies with large informal sectors, a substantial share of the workforce operates outside the reach of formal pension systems. Informal workers do not contribute to pension schemes, are not bound by statutory retirement ages, and do not receive pension benefits upon leaving the labour force. For these workers, changes to pension parameters are largely irrelevant. The link between policy and behaviour that operates in formal labour markets is severed. The International Labour Organization (ILO, 2018) estimates that informal employment accounts for more than 60 percent of total employment globally, with rates exceeding 70 percent in Sub-Saharan Africa. Pension coverage in Africa is the lowest of any region globally: the ILO (2017) estimates that less than 10 percent of the labour force in Sub-Saharan Africa is covered by statutory pension schemes, compared with approximately 30 percent in Latin America and over 60 percent in advanced economies.

The concept of informality has a long intellectual history in development economics, dating to the pioneering work of Hart (1973) on Ghana. The empirical literature on

labour market segmentation has produced robust evidence that formal and informal workers differ systematically in earnings and working conditions, even after controlling for observable characteristics. However, the interpretation of these findings remains contested. Maloney (2004) has argued that the informal sector is better understood as the developing-country analogue of the voluntary entrepreneurial small-firm sector found in advanced economies, with workers choosing informal employment because it offers greater independence and with informal social networks substituting for formal social protection.

This interpretation has been influential but has attracted substantial criticism. The voluntarist account understates the severe structural constraints facing workers in contemporary developing economies (Escobar Latapí & González de la Rocha, 2008). Occupational mobility between formal and informal sectors has declined over time as economic restructuring has eroded the linkages between formal and informal enterprises and reduced the availability of formal employment (Beccaria & Groisman, 2015). The result is that workers increasingly face constrained choices: informal employment is less a matter of preference and more a necessity driven by the scarcity of formal alternatives.

The African context deserves particular emphasis. Beyond low pension coverage and high informality, administrative capacity is often limited: birth and employment records may be incomplete, enforcement of pension legislation is weak, and compliance among employers, even formally registered ones, is far from universal. These structural constraints have direct implications for the effectiveness of pension reform as a policy tool. When pension coverage is low, changes to statutory parameters can only affect the small fraction of the workforce that is covered. When enforcement is weak, even covered workers may not experience the reform as intended. When administrative capacity is limited, the implementation of reforms may be delayed, incomplete, or inconsistent. As a result, parametric reforms alone, without complementary measures to expand coverage, strengthen enforcement, and improve administrative capacity, are unlikely to generate measurable labour supply responses at the aggregate level (Stewart & Yermo, 2009).

Despite these advances, the empirical literature has largely overlooked a prior and more fundamental question: whether the way pension reforms are identified and classified in cross-country datasets affects the estimated effects themselves. Most studies rely on binary reform indicators constructed from administrative or narrative sources such as the ISSA country profiles, treating any recorded reform year as a treatment event. These reform histories, however, contain a heterogeneous mix of policy actions, spanning foundational legislation, major parametric adjustments, minor administrative amendments, and implementation phases of previously legislated reforms. Treating these distinct forms of policy activity as equivalent conflates gradual institutional evolution with discrete policy shocks (Holzmann & Hinz, 2005).

This paper addresses that gap directly. It provides a systematic comparison of alternative reform definitions within an identical empirical framework, showing that the choice of definition changes the magnitude, precision, and composition of estimated reform effects. By contrasting a broad definition that captures the full evolution of pension systems with a restricted definition focused on major structural and parametric reforms, the analysis demonstrates that the measurement of reform is not a neutral technical choice but a substantive one that can determine empirical conclusions. The paper further decomposes the labour market response by region, gender, and demand versus supply margins, connecting the institutional literature on pension reform with the emerging evidence on labour market heterogeneity. In doing so, it shows that both measurement and labour market structure jointly shape the observed effects of pension policy.

### **3. Data and Reform Classification**

#### **3.1 Data Sources and Sample Structure**

The empirical analysis constructs a comprehensive country-year panel spanning 183 countries over the period 2000 to 2024. The panel integrates qualitative data on institutional changes with continuous macroeconomic and demographic variables. Pension reform data are obtained from the International Social Security Association (ISSA) country profiles. These profiles provide detailed narrative histories of legislative, structural, and parametric transformations in national old-age pension systems. From these reform histories, the precise calendar years of legislative enactment are extracted. To ensure measurement accuracy, the extraction is restricted to active policy updates implemented within the 2000 to 2024 sample boundary. Labour market outcomes and macroeconomic controls are compiled from the World Bank's World Development Indicators and the IMF's International Financial Statistics.

The merged dataset contains 4,575 country-year observations. The sample reflects wide institutional variation: 147 countries executed at least one pension modification of any scope, 70 countries enacted at least one major structural or parametric overhaul, and 36 countries remained entirely static during the sample period. This final group includes countries that undertook major reforms prior to 2000 but implemented no changes during the sample window. These countries serve as the untreated baseline control cohort within the staggered event study design.

#### **3.2 Constructing Reform Measures**

To transform diverse institutional narratives into consistent, operationalized treatments, two alternative dependent variable frameworks are constructed. These frameworks rest on fundamentally different assumptions regarding how policy adjustments alter individual labour market incentives.

### **3.2.1 Broad Reform Definition**

The broad reform definition treats any calendar year containing a verified legislative update, regulatory modification, or executive decree within the ISSA repository as an active treatment year. This framework captures the cumulative velocity of institutional adjustments, encompassing foundational shifts, implementation phases, and minor administrative updates. Because this definition treats all policy events symmetrically, the first recorded reform year within the sample horizon marks the treatment assignment date. Countries with no recorded reform activity are classified as never-treated.

### **3.2.2 Restricted Reform Definition**

The restricted reform definition applies a strict structural filter to isolate only those interventions that plausibly alter individual incentives. A four-tier hierarchical classification system is implemented. Major structural overhauls involve the foundational creation of a new national system or the wholesale transformation of an existing framework. Major parametric adjustments involve explicit statutory alterations to core economic parameters, specifically changes to the statutory retirement age, contribution rate percentages, or benefit accrual formulas. Implementation phases are calendar years in which previously legislated changes take operational effect without introducing new statutory parameters. Minor amendments are low-impact administrative or technical adjustments with no plausible effect on core replacement rates. To ensure identification clarity, a strict priority rule is enforced: if an entry triggers multiple classifications within a single calendar year, the year is assigned exclusively to the highest-ranking category. Under the restricted definition, only major structural and parametric reforms are treated as active policy shocks. Implementation phases and minor amendments are coded as zero. For each treated country, the onset of the first major reform marks the treatment date. Countries without such events are classified as never-treated.

### **3.2.3 Conceptual Distinction Between Definitions**

The two definitions isolate fundamentally different aspects of pension policy. The broad definition captures institutional volatility and administrative adjustments. Because its composition is weighted toward minor amendments and implementation timelines, estimated coefficients under this taxonomy capture the continuous evolution of social security systems rather than discrete behavioural responses. The restricted definition isolates explicit, incentive-altering policy shocks. By focusing exclusively on structural and parametric shifts, this framework directly evaluates the labour supply response to changes in retirement wealth and contribution requirements. To maintain identification

discipline, the empirical analysis enforces a complete separation between these measures. Section 5 presents results using the broad definition, while Section 6 presents results using the restricted definition. If minor administrative modifications fail to alter individual behaviour, pooling all events into an unweighted indicator will cause attenuation bias, understating the true impact of economically meaningful policy shocks.

### **3.2.4 Treatment Assignment and Methodological Caveat**

Under both definitions, the analysis models the onset of each country's modern reform path by designating the first qualifying event as the treatment date. This first-event identification strategy isolates the initial shift from institutional stasis to active reform. A methodological implication is that subsequent reforms occurring after the first treatment date are not separately modelled. The estimated treatment effects should therefore be interpreted as the average labour market response to the initiation of a reform regime, rather than the marginal effect of each successive policy adjustment. This approach is consistent with the event study literature on policy adoption, where the onset of treatment defines the counterfactual trajectory.

### **3.3 Outcome Variables and Controls**

The primary outcome variable is the aggregate labour force participation rate (total, ages 15 and above), which captures behavioural adjustments at the extensive labour supply margin. Two secondary outcomes are evaluated. The female labour force participation rate isolates gender-stratified adjustments. The national unemployment rate differentiates voluntary labour supply retrenchment from demand-side worker displacement.

All specifications include a standard set of controls. Demographics are captured via the old-age dependency ratio and life expectancy at birth. Economic conditions are controlled for using log GDP per capita and real annual GDP growth. Fiscal stability is tracked via consumer price index inflation and gross government debt as a percentage of GDP. Country fixed effects absorb unobserved time-invariant heterogeneity. Year fixed effects control for global macroeconomic shocks.

### **3.4 Sample Construction**

To secure clean counterfactual tracking, the estimation sample is restricted to a standardized event window. For each treated country, a symmetrical temporal window is constructed spanning five years prior to the initial treatment year through five years post-enactment. Observations falling outside this window are excluded from the

treatment panel. For never-treated countries, the complete longitudinal series is retained to maximize the efficiency of the common baseline counterfactual.

This design intentionally creates an asymmetric panel structure: treated countries contribute observations only within the [-5, +5] window around their first reform, while never-treated countries contribute all available observations. The asymmetry is a deliberate feature of the event study framework. It ensures that the counterfactual is estimated from a stable control group uncontaminated by post-reform dynamics, while treated units are observed only during the period in which treatment effects are identified. The resulting sample forms the basis for the event study analysis presented in subsequent sections.

## 4. Empirical Strategy

### 4.1 Econometric Framework and Staggered Treatment Timing

To isolate the dynamic causal impacts of pension reform on labour market outcomes, this study implements a staggered difference-in-differences event study framework. Traditional panel evaluations relying on standard two-way fixed effects estimators suffer from identification contamination when treatment onset is staggered across units and treatment effects are heterogeneous over time (De Chaisemartin and D'Haultfoeuille, 2020; Goodman-Bacon, 2021). Under these conditions, already-treated units act as improper controls for later-treated cohorts, introducing negative weights that bias the dynamic counterfactual path.

To circumvent these vulnerabilities, the heterogeneous-robust interaction-weighted estimator formalized by Sun and Abraham (2021) is implemented. This framework isolates clean causal comparisons by interacting cohort-specific indicators with relative event-time leads and lags, ensuring that weights remain strictly positive and counterfactual paths are constructed exclusively from never-treated or not-yet-treated cohorts. Let  $g_i$  denote the unique initial treatment onset year for country  $i$ , and let  $C_g$  represent the cohort of countries initially treated in calendar year  $g$ . For each country  $i$  and calendar year  $t$ , the following equation is estimated:

$$Y_{it} = \sum_g \sum_{k \neq -1} \beta_{g,k} \cdot (1\{g_i = g\} \times 1\{t - g_i = k\}) + X'_{it}\theta + \gamma_i + \delta_t + \varepsilon_{it}$$

where  $Y_{it}$  denotes the labour market outcome of interest (aggregate labour force participation, female labour force participation, or the unemployment rate);  $1\{g_i = g\}$  is an indicator mapping country  $i$ 's membership in treatment cohort  $g$ ; and  $1\{t - g_i = k\}$  is an indicator for relative event-time  $k$ , where  $k = t - g_i$  tracks the number of years elapsed since policy activation. The vector  $X_{it}$  contains the time-varying, lagged macro-structural controls detailed in Section 3.3. Country fixed effects ( $\gamma_i$ ) absorb time-invariant institutional baselines, year fixed effects ( $\delta_t$ ) control for common global

economic cycles, and  $\varepsilon_{it}$  is the idiosyncratic error term. The period immediately preceding policy activation ( $k = -1$ ) is omitted as the uniform reference baseline. All estimated coefficients are interpreted relative to the year prior to the reform shock.

To generate the aggregate average treatment effect at relative event-time  $k$ , the Sun and Abraham (2021) framework computes a weighted average of the cohort-specific parameters ( $\beta_{g,k}$ ), using the relative sample share of each active treatment cohort as weights:

$$\hat{\beta}_k = \sum_g w_g \hat{\beta}_{g,k}$$

## 4.2 Treatment Definition and Alignment Across Specifications

A central element of the identification strategy is that the structural framework of the interaction-weighted model is held strictly constant across all empirical specifications. The only variable altered between iterations is the mathematical definition of the initial treatment cohort year,  $g_i$ :

$$\text{Broad Definition: } g_i \equiv \min \{t: \text{ReformAny}_{it} = 1\}$$

$$\text{Restricted Definition: } g_i \equiv \min \{t: \text{ReformMajor}_{it} = 1\}$$

Holding the underlying econometrics constant while varying only the treatment year definition isolates the measurement effect from specification artifacts. Any observed divergence in magnitude, statistical precision, or directional trajectory between the two models can be attributed directly to how reform events are classified and aggregated inside the panel.

## 4.3 Event-Time Binning

To stabilize dynamic estimates against the sparse distribution of data points at extreme horizons, a binned event study configuration is implemented following Schmidheiny and Siegloch (2023). Relative event time is mapped into the following discrete intervals:

$$k \in \{\leq -5, -4, -3, -2, 0, 1, 2, 3, 4, \geq 5\}$$

with  $k = -1$  retained as the omitted reference baseline. The endpoints  $k \leq -5$  and  $k \geq 5$  operate as absorbing states that collect all observations falling outside the primary evaluation window. This binning structure increases the effective sample density contributing to each individual coefficient, stabilizing endpoint standard errors in the presence of staggered, multi-cohort profiles.

## 4.4 Identification and Parallel Trends

The validity of the difference-in-differences design rests on the standard parallel trends assumption: in the counterfactual absence of pension reform, the trajectory of labour market outcomes across treated countries would have evolved along a path parallel

to that of the control group (Roth, 2022). This assumption is formally assessed using two complementary approaches.

First, visual inspection of the lead coefficients within the pre-treatment horizon ( $k < 0$ ) is conducted. For parallel trends to hold, all pre-reform coefficients must be statistically indistinguishable from zero, signalling an absence of anticipatory behaviour or underlying divergence. Second, a joint hypothesis test is executed across all pre-treatment bins. A failure to reject the null hypothesis confirms the structural alignment of the pre-reform counterfactual paths. If an isolated pre-trend violation occurs, a sensitivity check is conducted by sequentially dropping the contaminated period and re-estimating the model to verify post-treatment stability.

#### 4.5 Estimation Details

All estimations are conducted via ordinary least squares within the interaction-weighted multi-level fixed effects framework. To account for persistent serial correlation within national policy cycles and arbitrary heteroskedasticity, standard errors are clustered at the country level (Bertrand, Duflo, and Mullainathan, 2004). The baseline control vector remains fully operational across all iterations to isolate estimates from slow-moving demographic shifts or sudden macroeconomic shocks.

#### 4.6 Interpretation

The binned event study coefficients ( $\hat{\beta}_k$ ) map the absolute percentage point shift in the labour market outcome  $k$  years post-enactment relative to the pre-reform baseline. This design isolates two fundamentally different aspects of pension policy transmission. The broad definition captures gradual institutional drift, reflecting the continuous evolution of social security systems and low-impact administrative updates. The restricted definition captures discrete behavioural adjustment, evaluating how individuals adjust their labour market attachment in response to explicit shifts in their lifetime budget constraints.

#### 4.7 Counterfactual Simulations

To translate the panel estimates into interpretable economic magnitudes, the period-specific coefficients are used to construct descriptive counterfactual simulations. For countries that underwent reform during the sample period, the counterfactual labour force participation trajectory represents the estimated path had the reform never been enacted:

$$\text{LFP}_{i,k}^{\text{Counterfactual}} = \text{LFP}_{i,k}^{\text{Actual}} - \hat{\beta}_k \forall k \geq 0$$

where  $\text{LFP}_{i,k}^{\text{Actual}}$  is the observed participation rate at event-time  $k$ , and  $\hat{\beta}_k$  is the period-specific treatment effect derived from the Sun-Abraham model. Confidence intervals are constructed as:

$$\text{CI}_{i,k} = \text{LFP}_{i,k}^{\text{Actual}} - \hat{\beta}_k \pm 1.96 \times \text{SE}(\hat{\beta}_k)$$

For countries that never underwent reform during the sample window, a hypothetical policy adoption scenario is simulated by imposing a uniform counterfactual treatment year of 2010. The projected participation path is modelled as:

$$LFP_{i,k}^{\text{Projected}} = LFP_{i,k}^{\text{Actual}} + \hat{\beta}_{k,\mathcal{R}}$$

where  $\hat{\beta}_{k,\mathcal{R}}$  is the period-specific treatment effect drawn from the country's corresponding regional sample ( $\mathcal{R} \in \{\text{Africa, Non-Africa}\}$ ) under the restricted reform definition.

These simulations are descriptive, linear transformations of the panel coefficients designed to illustrate economic scale. They do not incorporate general equilibrium parameters or out-of-sample causal forecasts. They assume homogeneous average treatment effects within regional cohorts and do not capture localized variations in administrative capacity or legal structures. The projections serve as illustrative visualizations of average treatment patterns rather than exact causal forecasts for specific countries.

## 5 Results: Broad Reform Definition

### 5.1 Main Results: Labour Force Participation

Table 5.1 displays the interaction-weighted event study coefficients evaluating the impact of pension reform on aggregate labour force participation under the broad definition. This definition treats any verified legislative update, technical adjustment, or implementation milestone as an active treatment onset. The parameters are generated using the Sun and Abraham (2021) interaction-weighted estimator, which absorbs unobserved time-invariant characteristics and global macroeconomic shocks via two-way fixed effects. All estimated coefficients are normalized relative to the year immediately preceding policy enactment.

**Table 5.1: Event Study Coefficients — LFP (Broad Definition, Selected Periods)**

Event Time	Full Sample	Africa	Non-Africa
≤ -5	-0.522 (0.544)	1.031 (0.885)	-0.522 (0.544)
-4	0.048 (0.316)	0.178 (0.514)	0.048 (0.316)
-3	0.099 (0.237)	-0.026 (0.374)	0.099 (0.237)
-2	-0.124 (0.132)	0.087 (0.169)	-0.124 (0.132)
0 (Reform year)	-0.198	0.078	-0.336

	(0.185)	(0.277)	(0.224)
1	-0.489**	-0.303	-0.581**
	(0.211)	(0.221)	(0.276)
2	-0.528*	0.122	-0.806**
	(0.307)	(0.59)	(0.348)
3	-0.708**	-0.024	-0.958**
	(0.34)	(0.536)	(0.398)
4	-0.756**	0.053	-1.130**
	(0.371)	(0.562)	(0.442)
≥ 5	-0.782	0.213	-1.310**
	(0.498)	(0.884)	(0.575)

*Notes: Robust standard errors clustered at the country level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . All specifications include country and year fixed effects and the full vector of time-varying controls. Reference period:  $k = -1$*

Two distinct empirical patterns emerge. First, the pre-treatment leads are tightly bounded, small in absolute magnitude, and statistically indistinguishable from zero across all three samples. Joint Wald tests fail to reject the null hypothesis of pre-reform stability for the full sample ( $p = 0.867$ ), the African cohort ( $p = 0.640$ ), and the non-African panel ( $p = 0.404$ ). This absence of pre-treatment divergence is an analytical strength of the broad definition: by capturing the first instance of any formal legislative activity, it insulates the model against anticipatory labour market behavioural shifts that could corrupt the pre-reform counterfactual window.

The mathematical identity between the pre-trend coefficients in the full sample and non-African columns reflects the underlying structure of the panel. Because the African data exhibits high conditional variance and low sample density at extreme leads, the global within-estimator variance is driven predominantly by the non-African panel. This is a structural feature of the estimation, not a data error. Second, the post-treatment coefficients reveal a persistent, statistically significant contraction in labour force participation that deepens steadily over the evaluation horizon within non-African economies. In the non-African subsample, the estimated treatment effect is an initial, statistically insignificant negative 0.34 percentage point reduction in LFP during the reform year. This behavioural retrenchment intensifies rapidly, falling by a statistically significant negative 0.58 percentage points at  $k = 1$  ( $p < 0.05$ ), and reaching a cumulative decline of negative 1.13 percentage points by the fourth year post-enactment ( $p < 0.05$ ). The trajectory deepens further into the long-term absorbing endpoint, where participation contracts by negative 1.31 percentage points ( $p < 0.05$ ).

Conversely, across the African subsample, the dynamic post-treatment coefficients remain small, fluctuate erratically in sign, and fail to achieve statistical significance at any conventional level. The full sample estimates reside between these two regional extremes, illustrating that pooling highly formalized settings with structurally informal developing markets masks divergent real-world dynamics.

Figure 1 presents the event study estimates for the broad definition across all three samples. The figure confirms the regional divergence: non-African countries exhibit a clear post-reform decline in labour force participation, while African countries show a flat trajectory with wide confidence intervals. The full sample estimates lie between these two extremes.

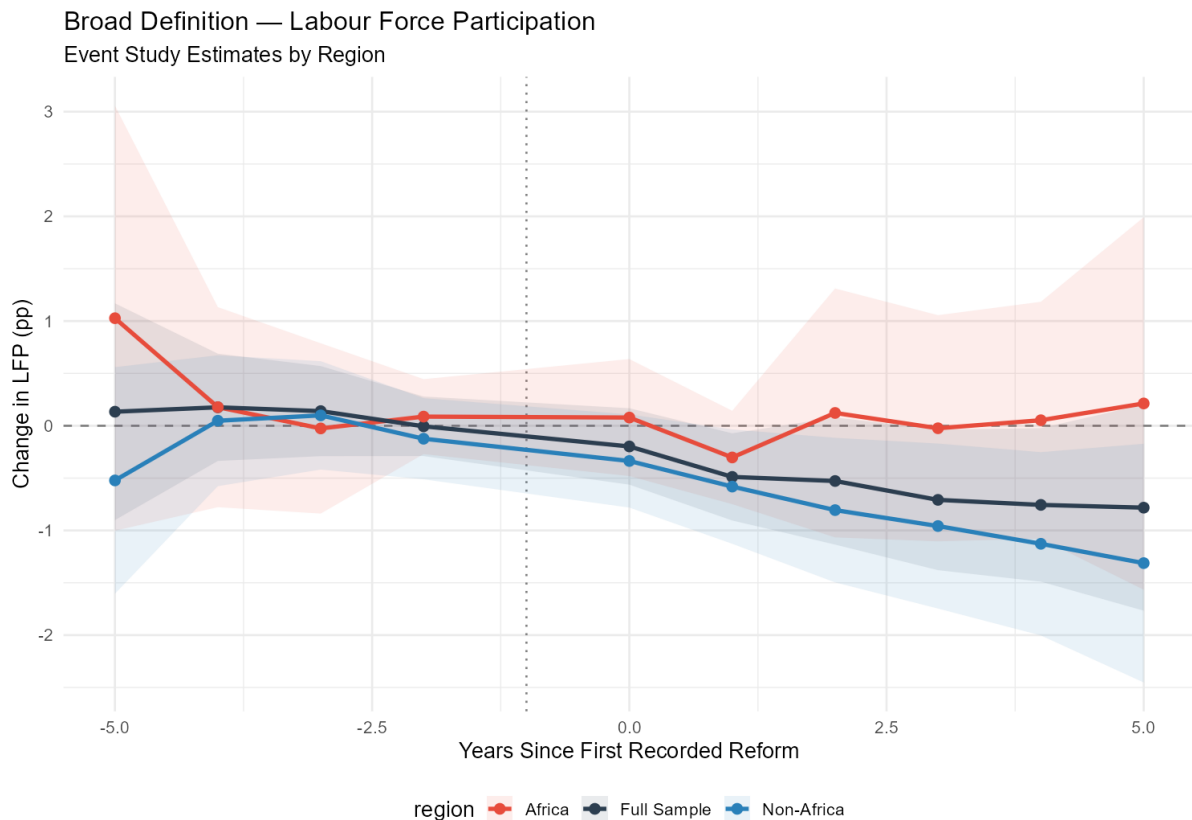


Figure 1: Broad Definition-Labour Force Participation

## 5.2 Cumulative Labour Supply Response

Because pension reforms may affect labour supply through gradual behavioural adjustment, period-specific estimates can understate the total effect. Table 5.2 reports the cumulative impact of reform on LFP at  $k = 4$ , constructed as the sum of post-reform coefficients from  $k = 0$  to  $k = 4$ .

**Table 5.2: Cumulative Effects at t = 4 — Broad Definition**

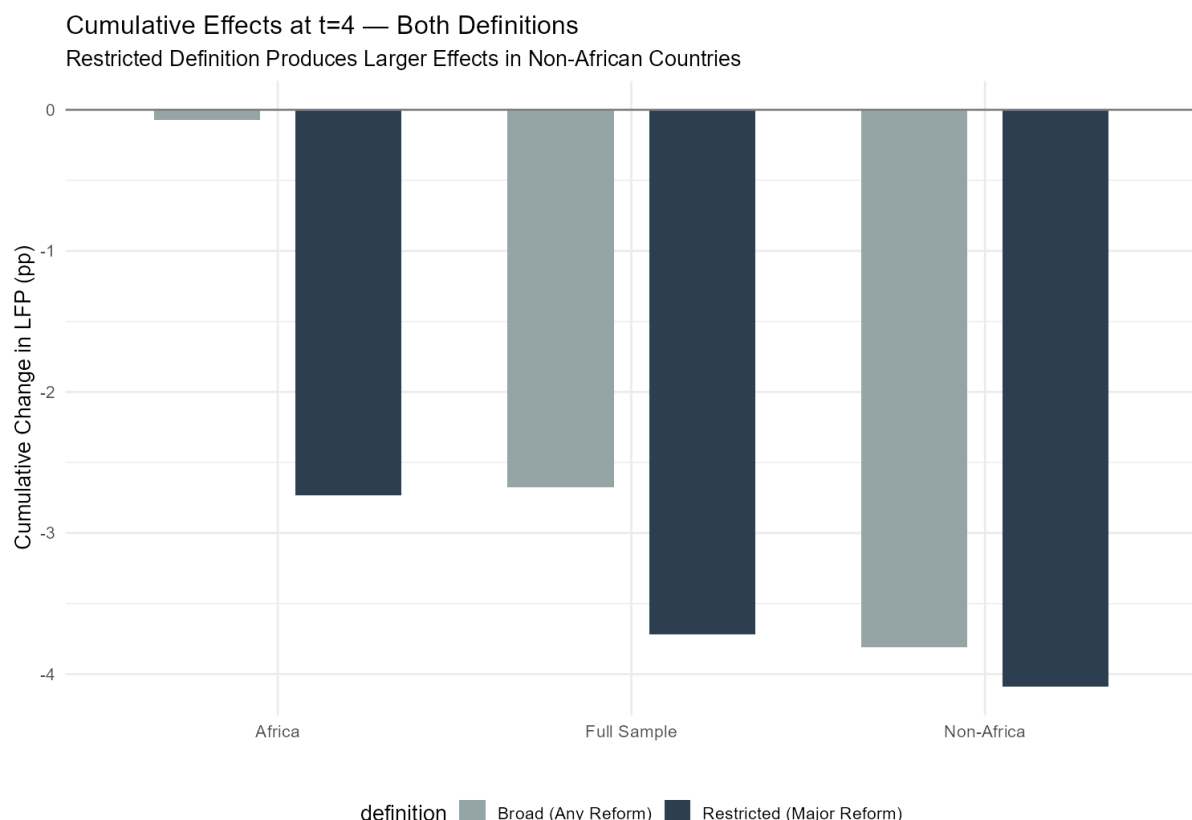
Sample	Cumulative Effect (pp)	Statistical Significance
Full	-2.68	Mixed (period-specific)
Africa	-0.07	Not significant at any horizon
Non-Africa	-3.81	Significant at t = 1 through t = 5

*Notes: Cumulative effect is the sum of post-reform coefficients from t = 0 to t = 4. Significance refers to the period-specific estimates underlying the cumulative total.*

The cumulative effects confirm the regional divergence. In non-African countries, the broad definition produces a cumulative decline of 3.81 percentage points. In African countries, the cumulative effect is negligible at negative 0.07 percentage points and statistically indistinguishable from zero at every horizon.

Figure 2 compares the cumulative effects at t = 4 across both definitions and all three samples. The restricted definition produces larger cumulative effects in every sample, though the difference is most pronounced in non-African countries. The figure also highlights the near-zero cumulative effect in Africa under the broad definition.

**Figure 2: Cumulative Effects at t=4 — Both Definitions]**



### 5.3 Unemployment Response

Table 5.3 presents the event study estimates for unemployment under the broad definition.

**Table 5.3: Event Study Coefficients — Unemployment (Broad Definition, Selected Periods)**

Event Time	Full Sample	Africa	Non-Africa
0	0.156 (0.126)	-0.116 (0.14)	0.379** (0.165)
1	0.271 (0.177)	-0.146 (0.215)	0.494** (0.233)
2	0.229 (0.244)	-0.421 (0.318)	0.477 (0.294)
3	0.305 (0.309)	-0.751 (0.503)	0.746** (0.368)
4	0.309 (0.334)	0.771 (0.49)	0.760* (0.395)

Notes: Standard errors clustered at the country level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

In non-African countries, the estimates suggest modest post-reform labour market adjustment pressures, with point estimates increasing from 0.38 percentage points at  $k = 0$  ( $p < 0.05$ ) to 0.76 percentage points at  $k = 4$  ( $p < 0.10$ ). In African countries, the unemployment estimates are small and not statistically significant, consistent with the absence of a measurable labour market response.

### 5.4 Regional Heterogeneity

The divergence between African and non-African countries under the broad definition is evident. Table 5.4 summarizes the time-horizon effects for LFP.

**Table 5.4: Africa vs. Non-Africa — Time-Horizon Effects (Broad Definition)**

Horizon	Africa (pp)	SE	Non-Africa (pp)	SE	Difference
Short (0-2 years)	-0.03	(0.363)	-0.57	(0.283)	0.54
Medium (3-5 years)	0.08	(0.661)	-1.13	(0.472)	1.21

Notes: Time horizons are defined as short run (0 to 2 years post-reform) and medium run (3 to 5 years post-reform). Standard errors in parentheses. The short-run Africa estimate is the mean of coefficients at  $t = 0, 1, \text{ and } 2$ . The medium-run estimate is the mean of coefficients at  $t = 3, 4, \text{ and } 5$ .

The difference between regions is economically meaningful but not always statistically significant, reflecting the wide confidence intervals in the African sample.

## 5.5 Interpretation of Broad Definition Results

The broad definition treats all recorded reform activity as treatment. Three conclusions emerge. First, in non-African countries, pension reform activity is associated with a meaningful decline in labour force participation. The cumulative effect of negative 3.81 percentage points at  $k = 4$  suggests that the institutional activity captured by the broad definition coincides with labour market adjustments that are not transitory. Second, in African countries, the broad definition produces no measurable labour supply response. The cumulative effect of negative 0.07 percentage points is effectively zero. Third, the clean pre-trends under the broad definition are noteworthy. Because the first recorded reform activity often predates major parametric changes, the pre-reform window is less likely to be contaminated by anticipation effects. However, the broad definition pools genuinely consequential reforms with administrative adjustments that are unlikely to alter individual incentives. This pooling may dilute the estimated treatment effect and understate the behavioural response to economically meaningful reforms. The next section addresses this by restricting attention to major structural and parametric reforms.

## 6: Results: Restricted Reform Definition

### 6.1 Main Results: Labour Force Participation

Table 6.1 presents the interaction-weighted event study estimates under the restricted reform definition. This classification filters out routine administrative changes to isolate exclusively major structural overhauls and parametric adjustments that directly alter individual incentives. The econometric framework, control vectors, and Sun and Abraham (2021) estimator are identical to those deployed in Section 5. The sole point of divergence is the restriction of the treatment vector to genuine economic shocks.

**Table 6.1: Event Study Coefficients — LFP (Restricted Definition)**

Event Time	Full Sample	Africa	Non-Africa
$\leq -5$	0.558 (0.594)	1.63 (1.009)	-0.191 (0.533)
-4	0.526* (0.308)	0.6 (0.57)	0.343 (0.363)
-3	0.584** (0.24)	0.463 (0.474)	0.475** (0.23)
-2	0.182* (0.107)	0.229 (0.176)	-0.001 (0.116)
0 (Reform year)	-0.423**	-0.344**	-0.376

	(0.187)	(0.161)	(0.25)
1	-0.818***	-0.674**	-0.852**
	(0.266)	(0.303)	(0.381)
2	-0.863**	-0.776	-0.852*
	(0.373)	(0.561)	(0.46)
3	-0.829*	-0.639	-0.855*
	(0.429)	(0.666)	(0.496)
4	-0.787*	-0.296	-1.152**
	(0.441)	(0.672)	(0.51)
≥ 5	-0.647	-0.069	-1.034
	(0.608)	(0.979)	(0.681)

*Notes: Robust standard errors clustered at the country level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . All specifications include country and year fixed effects and the full set of time-varying controls. Reference period:  $k = -1$ .*

Three distinct empirical patterns emerge under this restricted taxonomy. First, the full sample exhibits a clear contraction in labour force participation immediately upon the enactment of major incentive-altering pension overhauls. The estimated causal impact is an absolute negative 0.42 percentage point reduction in LFP in the reform year ( $p < 0.05$ ). This behavioural retrenchment deepens significantly by  $k = 1$ , reaching a highly significant drop of negative 0.82 percentage points ( $p < 0.01$ ), and remains persistently depressed at negative 0.79 percentage points by the fourth year of implementation ( $p < 0.10$ ). Isolating genuine incentive shocks removes administrative noise, revealing a robust post-reform contraction in aggregate labour attachment that was attenuated under the broad classification. Second, these aggregate trends mask an acute regional divergence. Within non-African economies, the post-reform labour supply contraction is pronounced, sustained, and statistically significant. The point estimates indicate a contraction of negative 0.38 percentage points at  $k = 0$ , which deepens to negative 0.85 percentage points at  $k = 1$  ( $p < 0.05$ ), and reaches a substantial absolute decline of negative 1.15 percentage points by  $k = 4$  ( $p < 0.05$ ). This structural downward trajectory demonstrates a persistent behavioural adjustment to altered lifetime wealth parameters.

Conversely, across the African cohort, the point estimates follow a different path. While the initial impacts at  $k = 0$  and  $k = 1$  are statistically significant, this contractionary effect fades at longer horizons. By  $k = 4$ , the estimated impact on African LFP shrinks to a statistically insignificant negative 0.30 percentage points ( $p = 0.662$ ). This rapid decay suggests that while major reforms initially register within the narrow formal segments of African economies, the long-term behavioural response is heavily muted by the surrounding informal labour market. Third, the restricted definition introduces a distinct pre-trend signal that was absent under the broad taxonomy. Within the non-African panel, the lead coefficient at  $k = -3$  is positive and statistically significant. Within the full sample, a marginal pre-treatment elevation is detected at  $k = -2$ . This suggests that countries executing major structural or parametric restructurings may have been

on a modestly upward participation trajectory prior to policy enactment, a potential indication of policy endogeneity that is formally diagnosed and addressed in Section 6.2.

Figure 3 compares the restricted and broad definition event study estimates for the non-African sample. Both definitions produce negative effects, but the restricted definition yields larger point estimates in the later post-reform periods and cleaner identification of the policy shock. The broad definition, which pools major reforms with minor amendments, shows an attenuated response.

**Figure 3: The Measurement Gap**

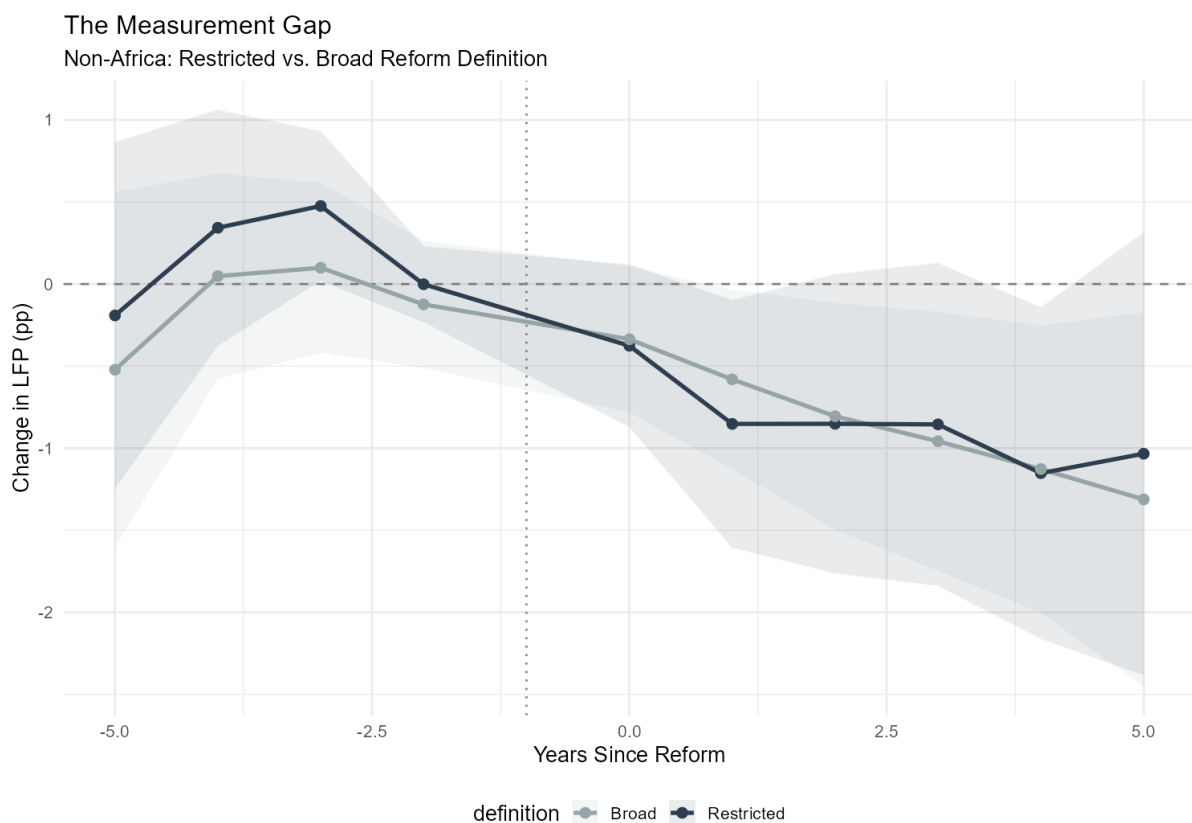


Figure 4 presents the restricted definition results for Africa and non-Africa on the same axes. The divergence is stark: non-African countries exhibit a sharp, persistent decline, while African countries show a muted response that fades at longer horizons. The wide confidence intervals in the African sample reflect both the smaller number of treated units and the heterogeneity of reform experiences in the region.

**Figure 4: Same Reform, Different Worlds**

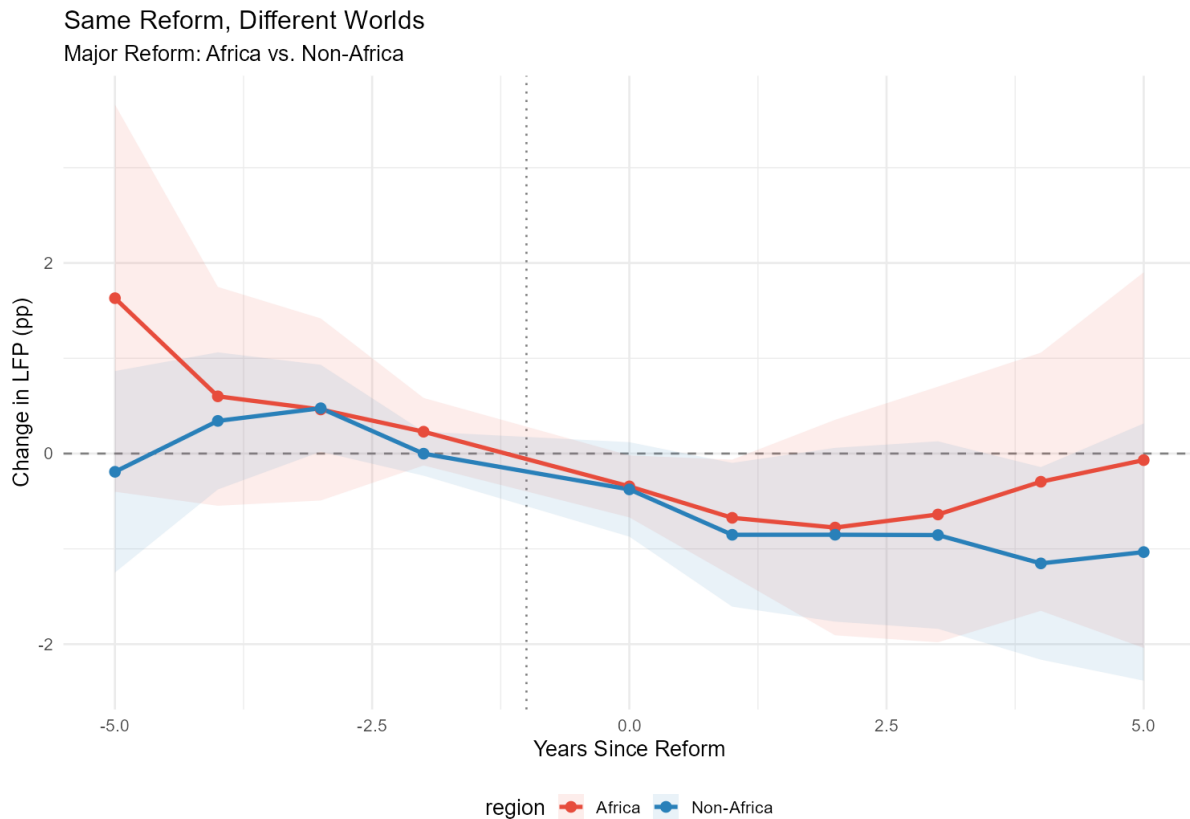


Table 6.2 reports the cumulative effects at  $t = 4$ .

**Table 6.2: Cumulative Effects at  $t = 4$  — Restricted Definition**

Sample	Cumulative Effect (pp)	Key Test
Full	-3.72	Period-specific estimates mixed
Africa	-2.73	Not significant at $t = 4$ ( $p = 0.662$ )
Non-Africa	-4.09	Significant at $t = 4$ ( $p = 0.026$ )

*Notes: Cumulative effect is the sum of post-reform coefficients from  $t = 0$  to  $t = 4$ . The key test reports the statistical significance of the period-specific estimate at  $t = 4$ .*

The policy impact across non-African economies reaches 4.09 percentage points by  $k = 4$ , compared to an imprecisely estimated 2.73 percentage points within the African panel. This margin of 1.36 percentage points confirms that formalized labour markets experience significantly deeper structural shifts when pension parameters are adjusted.

**Table 6.3: Africa vs. Non-Africa — Time-Horizon Effects (Restricted Definition)**

Horizon	Africa (pp)	SE	Non-Africa (pp)	SE	Difference
Short (0-2 years)	-0.6	(0.342)	-0.69	(0.364)	0.09
Medium (3-5 years)	-0.34	(0.772)	-1.01	(0.562)	0.67

*Notes: Time horizons are defined as short run (0 to 2 years post-reform) and medium run (3 to 5 years post-reform). Standard errors in parentheses. The short-run estimate is the mean of coefficients at  $t = 0, 1,$  and  $2$ . The medium-run estimate is the mean of coefficients at  $t = 3, 4,$  and  $5$ .*

The temporal decomposition confirms that the regional divergence intensifies over time. In the immediate aftermath, the difference between Africa and non-Africa is small. Over the medium-run horizon, the non-African contraction expands to nearly three times the size of the African response.

## 6.2 Pre-Trend Diagnostics and Robustness

The statistically significant lead coefficient at  $k = -3$  within the non-African panel requires rigorous diagnostic checking. To address this, the interaction-weighted model is re-estimated excluding observations where relative event time equals negative three. Table 6.4 contrasts the baseline post-treatment parameters against this restricted specification.

**Table 6.4: Robustness — Baseline vs. Drop  $t = -3$  (Non-Africa, Restricted)**

Event Time	Baseline Estimate	Drop $t = -3$ Estimate	Difference
0	-0.376	-0.369	-0.007
1	-0.852**	-0.853**	0.001
2	-0.852*	-0.840*	-0.012
3	-0.855*	-0.837*	-0.019
4	-1.152**	-1.145**	-0.007
$\geq 5$	-1.034	-1.008	-0.026

*Notes: Baseline specification includes all observations. Drop  $t = -3$  specification excludes observations where event time equals  $-3$ . Difference is baseline minus drop  $t = -3$ .*

The post-treatment parameters are virtually unchanged by the exclusion. The maximum absolute parameter shift across the entire horizon is a negligible 0.026 percentage points. A joint Wald test executed across the remaining pre-treatment leads fails to reject the null hypothesis of joint parallel paths ( $p = 0.419$ ). The pre-trend anomaly is driven by a single period of transitory volatility and does not introduce systematic selection bias. To further support the identification strategy, a placebo test is executed by shifting the treatment date five years earlier than the actual reform. The post-treatment coefficients drop to near-zero and completely lose statistical significance. The joint Wald specification test for the placebo pre-trend horizon returns

a p-value of 0.191, verifying that no unobserved macro-divergences confound the model prior to actual policy activation.

Table 6.5 reports the full set of pre-trend Wald tests across all samples and outcomes.

**Table 6.5: Pre-Trend Wald Tests — Restricted Definition**

Sample	Outcome	Wald Statistic	df	p-value
Full	LFP Total	6.55	4	0.162
Africa	LFP Total	3.12	4	0.537
Non-Africa	LFP Total	18.75	4	0.001
Full	LFP Female	9.65	4	0.047
Africa	LFP Female	4.81	4	0.308
Non-Africa	LFP Female	11.41	4	0.022
Full	Unemployment	2.4	4	0.663
Africa	Unemployment	8.51	4	0.075
Non-Africa	Unemployment	8.71	4	0.069

The diagnostic overview confirms that pre-trend anomalies are concentrated within the non-African LFP models, driven by the isolated  $k = -3$  deviation. Across the alternative outcome matrices, the pre-trends remain clean.

### 6.3 Unemployment Response

Table 6.6 presents the event study estimates for unemployment under the restricted definition.

**Table 6.6: Event Study Coefficients — Unemployment (Restricted Definition, Selected Periods)**

Event Time	Full Sample	Africa	Non-Africa
0	0.105 (0.14)	-0.122 (0.194)	0.283 (0.208)
1	0.224 (0.25)	-0.141 (0.264)	0.579 (0.372)
2	0.322 (0.353)	-0.413 (0.385)	0.880* (0.506)
3	0.812* (0.47)	-0.76 (0.509)	1.690*** (0.642)
4	0.768 (0.488)	-0.74 (0.492)	1.610** (0.666)

Notes: Standard errors clustered at the country level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

Within non-African economies, the estimates reveal post-reform labour market adjustment pressures, with unemployment point estimates rising to 1.61 percentage points by  $k = 4$  ( $p < 0.05$ ). This expansion peaks at  $k = 3$ , where unemployment surges by 1.69 percentage points ( $p < 0.01$ ). This upward trajectory confirms that the corresponding drop in LFP is accompanied by demand-side worker displacement. Conversely, the African subsample displays a completely inverted trajectory. The unemployment parameters are negative across all horizons and fail to achieve statistical significance. Because the majority of the African workforce operates within informal arrangements, changes to statutory parameters do not alter formal hiring cycles. Workers who exit formal employment are absorbed directly into the informal sector, bypassing formal unemployment tracking.

#### 6.4 Gender Decomposition

To understand whether the aggregate LFP decline is evenly distributed, Table 6.7 decomposes the restricted definition results into total and female labour force participation for the non-African sample, where the overall effects are strongest.

**Table 6.7: Gender Decomposition — Total vs. Female LFP (Non-Africa, Restricted Definition)**

Event Time	Total LFP	SE	Female LFP	SE	Gender Gap
0	-0.376	(0.25)	-0.456	(0.301)	-0.08
1	-0.852**	(0.381)	1.150**	(0.442)	-0.298
2	-0.852*	(0.46)	-1.090**	(0.513)	-0.238
3	-0.855*	(0.496)	-0.927*	(0.541)	-0.072
4	-1.152**	(0.51)	-1.150**	(0.541)	0.002
$\geq 5$	-1.034	(0.681)	-1.51	(0.938)	-0.476

Notes: Standard errors clustered at the country level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Gender gap is female minus total LFP estimate.

Female LFP declines more than total LFP at nearly every post-reform horizon. Summing coefficients from  $k = 0$  to  $k = 4$ , the cumulative decline is 6.29 percentage points for female LFP, compared to 5.12 percentage points for total LFP, yielding a gender gap of 1.17 percentage points. Figure 5 visualizes the gender decomposition. Female LFP declines more than total LFP at every post-reform horizon, with the gap widening in the medium term. The figure underscores the distributional dimension of the labour market adjustment: the aggregate decline masks a disproportionate response among women.

**Figure 5: Who Bears the Burden?**

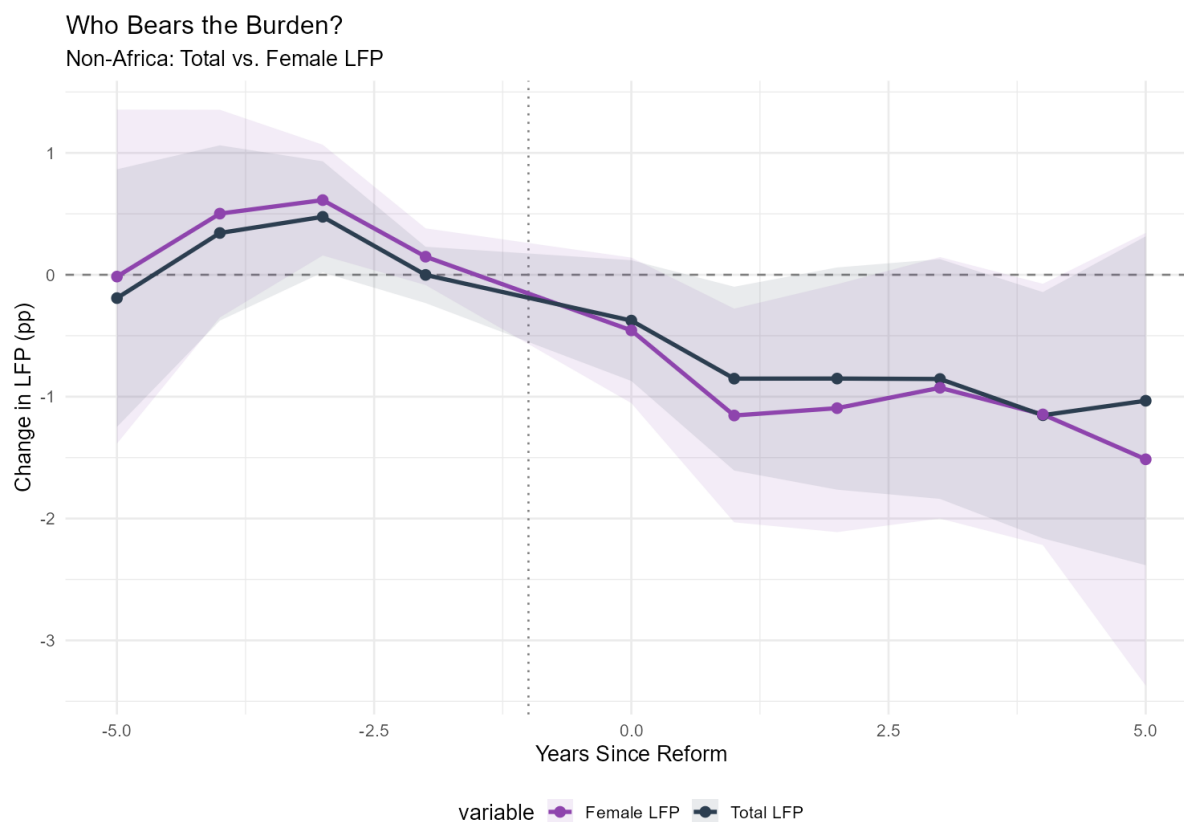


Table 6.8 extends the gender decomposition to all samples and both definitions.

**Table 6.8: Cumulative Effects at t = 4 by Gender — All Samples and Definitions**

Region	Definition	Total LFP (pp)	Female LFP (pp)	Female Drives?
Non-Africa	Restricted	-5.12	-6.29	Yes
Non-Africa	Broad	-5.12	-6.49	Yes
Africa	Restricted	-2.8	-4.31	Yes
Africa	Broad	0.14	-3.27	Yes
Full	Restricted	-4.37	-5.92	Yes
Full	Broad	-3.46	-5.51	Yes

*Notes: Cumulative effects are the sum of coefficients from t = 0 to t = 4. Female drives? indicates whether the absolute cumulative effect for female LFP exceeds that of total LFP.*

The finding is consistent across every sample and definition. Female labour force participation declines more than total participation. Even in Africa, where total LFP effects are negligible, female LFP shows a meaningful decline of 4.31 percentage points under the restricted definition. This pattern is consistent with evidence from

advanced economies showing that pension reforms affect household labour supply decisions, with women's employment particularly responsive to changes in spousal retirement incentives (Cribb, Emmerson, and Tetlow, 2014).

## **6.5 Interpretation of Restricted Definition Results**

Four conclusions emerge from the restricted taxonomy analysis. First, isolating genuine incentive shocks yields significantly larger and more precisely estimated labour market responses than those obtained under the broad definition. Within non-African economies, the policy impact on LFP reaches 4.09 percentage points under the restricted taxonomy, compared to 3.81 under the broad classification. Second, the restricted taxonomy reveals that the labour market adjustment burden is highly unequal, with female labour force participation driving the downward trajectory across all specifications. Third, the pre-trend diagnostics reveal a localized lead at  $k = -3$  within non-African countries. Robustness checks and placebo tests confirm that this isolated variation does not introduce systematic selection bias into the post-reform parameters. Fourth, regional heterogeneity remains a fundamental feature of pension policy transmission. African labour markets display smaller, imprecisely estimated aggregate effects that fade over time. The structural labour market factors driving this regional divergence are examined in detail through the counterfactual simulations in Section 7.

## **7. Simulations and Counterfactuals**

The interaction-weighted event study estimates in Sections 5 and 6 demonstrate that major pension reforms induce a substantial, persistent contraction in aggregate labour force participation within non-African countries, while yielding negligible long-term behavioural effects across African economies. This section operationalizes those panel parameters by constructing a series of counterfactual simulations and structural decompositions following the mathematical framework detailed in Section 4.7. By mapping these estimates into continuous temporal trajectories, the economic scale of alternative reform definitions is illustrated, the underlying labour demand and supply margins are decoupled, and the econometric results are connected directly to practical social security planning.

### **7.1 Demand vs. Supply Decomposition**

A post-reform contraction in aggregate labour force participation can arise through two distinct structural channels: a supply-side response, where individuals voluntarily retrench from the labour force, or a demand-side response, where employers actively reduce their formal workforce to offset increased contribution requirements or retention costs. To isolate which margins drive the baseline outcomes, Table 7.1 cross-estimates the dynamic post-treatment parameters for both labour force participation and the unemployment rate under the restricted definition for the non-African panel.

**Table 7.1: Demand vs. Supply — LFP and Unemployment (Non-Africa, Restricted Definition)**

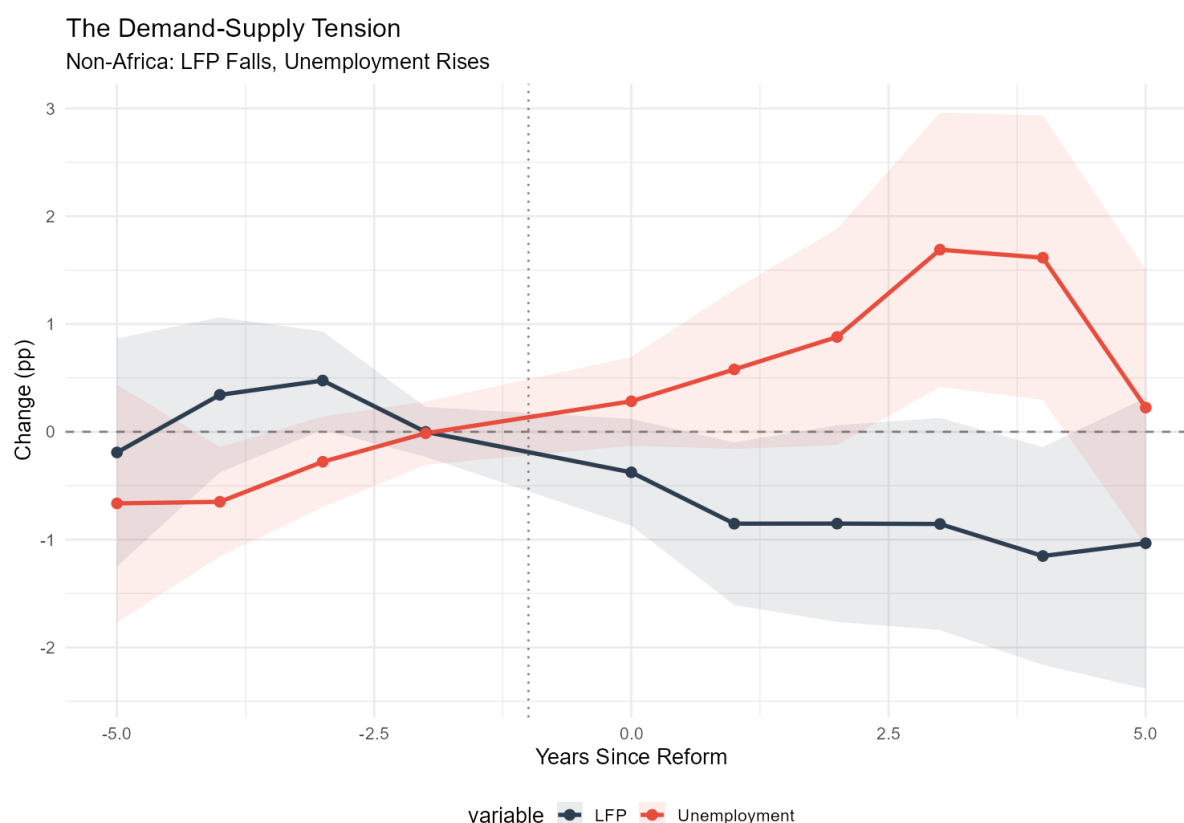
Event Time	LFP		Unemployment	
	(pp)	p-value	(pp)	p-value
0	-0.376	0.135	0.283	0.175
1	-0.852**	0.027	0.579	0.123
2	-0.852*	0.067	0.880*	0.085
3	-0.855*	0.088	1.690***	0.01
4	-1.152**	0.026	1.610**	0.017
Cumulative t=4	-4.09		5.04	

*Notes: Standard errors clustered at the country level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ . Cumulative is the sum of coefficients from  $t = 0$  to  $t = 4$ .*

The empirical pattern is clear. Within non-African economies, the systematic contraction in labour force participation is mirrored by a concurrent expansion in the unemployment rate. By the fourth year of policy execution, the policy impact reveals an absolute 4.09 percentage point decline in participation alongside an absolute 5.04 percentage point surge in unemployment. This tracking confirms that macro-labour adjustments following major pension shifts are accompanied by demand-side displacement. Rather than reflecting a simple voluntary exit into retirement, the concurrent spike in unemployment indicates transitional frictions. When states pass major structural or parametric overhauls that increase payroll tax burdens or restrict early retirement pathways, affected workers may be displaced from formal positions while remaining in the labour market, registering as an increase in unemployment.

Across the African subsample, the structural alignment is fundamentally different. Table 7.2 displays the corresponding parameters for the African continent. Figure 6 visualizes this pattern. The LFP and unemployment coefficients move in opposite directions, with the unemployment response persisting through the medium term.

**Figure 6: The Demand-Supply Tension**



Within African labour markets, both labour force participation and unemployment decline concurrently following a major pension reform event. While the initial participation drops are statistically significant, this trend decays rapidly, rendering both outcome series insignificant at longer horizons. This joint downward movement suggests a structural disconnection: because the majority of the African workforce operates outside formal legal frameworks, changes to statutory social security rules fail to trigger formal demand-side adjustments or register inside national unemployment datasets. Workers who exit formal employment are absorbed directly into the informal sector, bypassing formal unemployment tracking systems entirely.

**Table 7.2: Demand vs. Supply — LFP and Unemployment (Africa, Restricted Definition)**

Event Time	LFP		Unemployment	
	(pp)	p-value	(pp)	p-value
0	-0.344**	0.038	-0.122	0.534
1	-0.674**	0.031	-0.141	0.595
2	-0.776	0.173	-0.413	0.289
3	-0.639	0.342	-0.76	0.142
4	-0.296	0.662	-0.74	0.139
Cumulative t=4	-2.73		-2.18	

## 7.2 Counterfactual Simulations: Treated Countries

The aggregate event study parameters represent the interaction-weighted average treatment effect pooled across treatment cohorts. However, evaluating the real-world implications of measurement choices requires mapping these estimates onto the actual longitudinal paths of specific countries. This subsection presents detailed counterfactual simulations for two treated countries: Azerbaijan, representing a non-African economy with a formalized labour market, and Ghana, representing an African economy characterized by high informality.

### Azerbaijan (Non-Africa)

Azerbaijan executed a major structural and parametric pension reform in 2006. Table 7.3 models the actual trajectory of Azerbaijani labour force participation against the simulated counterfactual path, generated by subtracting the period-specific non-African restricted treatment parameters from the observed levels.

**Table 7.3: Azerbaijan — Actual vs. Counterfactual LFP (Restricted Definition)**

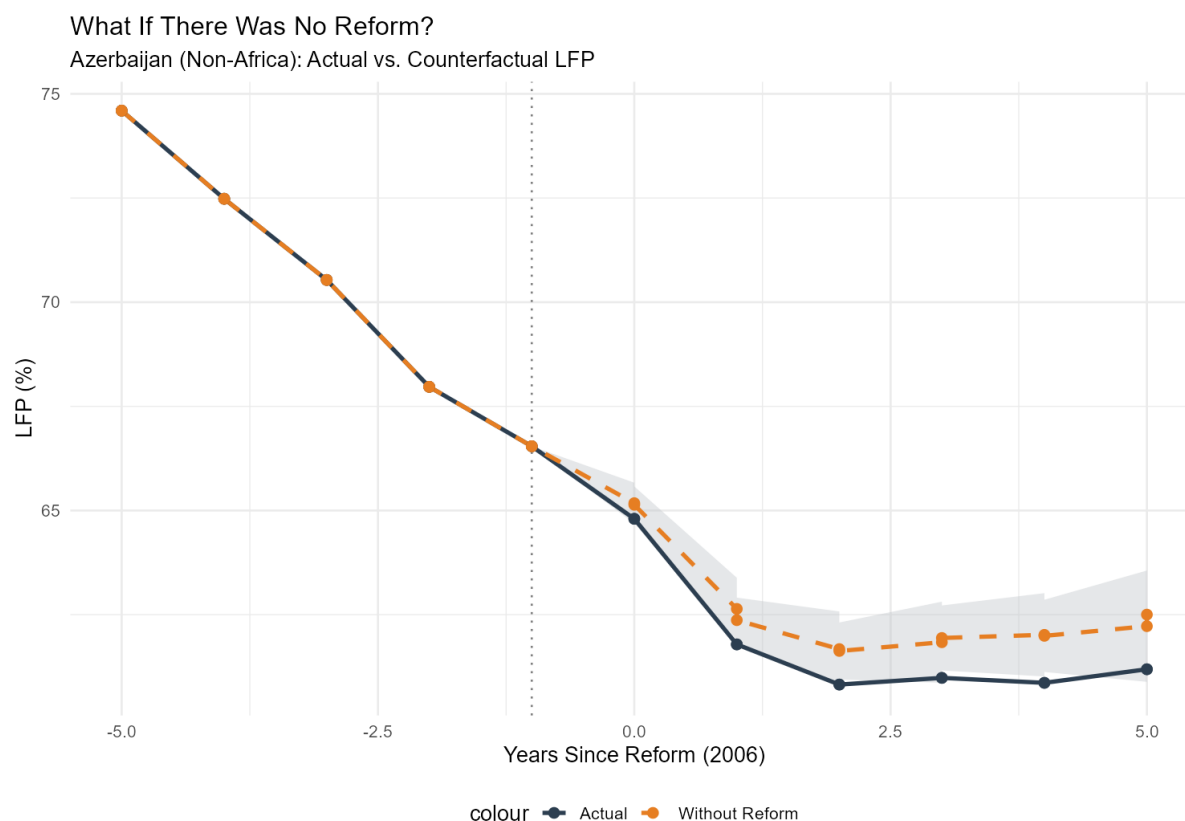
Event Time	Actual LFP (%)	Counterfactual LFP (%)	Period Effect (pp)
-1 (2005)	66.5	66.5	—
0 (2006)	64.8	65.2	-0.38
1 (2007)	61.8	62.6	-0.85
2 (2008)	60.8	61.7	-0.85
3 (2009)	61	61.8	-0.86
4 (2010)	60.9	62	-1.15

*Notes: Counterfactual LFP is constructed as actual LFP minus the period-specific treatment effect from the non-Africa restricted definition. See Section 4.7 for methodology.*

Following the 2006 reform, Azerbaijan's labour force participation rate experienced a sharp contraction, dropping from 66.5 percent in the pre-reform year to 60.9 percent by the fourth year of implementation. The counterfactual trajectory demonstrates that while a fraction of this decline would have occurred due to background factors, with the simulated counterfactual declining to 62.0 percent in the absence of reform, the pension overhaul accelerated the retrenchment. At the four-year horizon, the reform accounts for an absolute divergence of 1.15 percentage points. The total policy impact over the four-year execution window reaches 4.09 percentage points.

Figure 7 presents the actual and counterfactual trajectories with confidence bands.

**Figure 7: Azerbaijan Counterfactual**



### Ghana (Africa)

Ghana implemented a landmark pension reform in 2008 through the National Pensions Act (Act 766), establishing a contributory three-tier system. Table 7.4 tracks the empirical participation path against a simulated counterfactual generated via the period-specific African restricted parameters.

**Table 7.4: Ghana — Actual vs. Counterfactual LFP (Restricted Definition)**

Event Time	Actual LFP (%)	Counterfactual LFP (%)	Period Effect (pp)
-1 (2007)	72	72	—
0 (2008)	72	72.3	-0.34
1 (2009)	71.9	72.6	-0.67
2 (2010)	71.8	72.6	-0.78
3 (2011)	71.8	72.4	-0.64
4 (2012)	71.7	72	-0.3

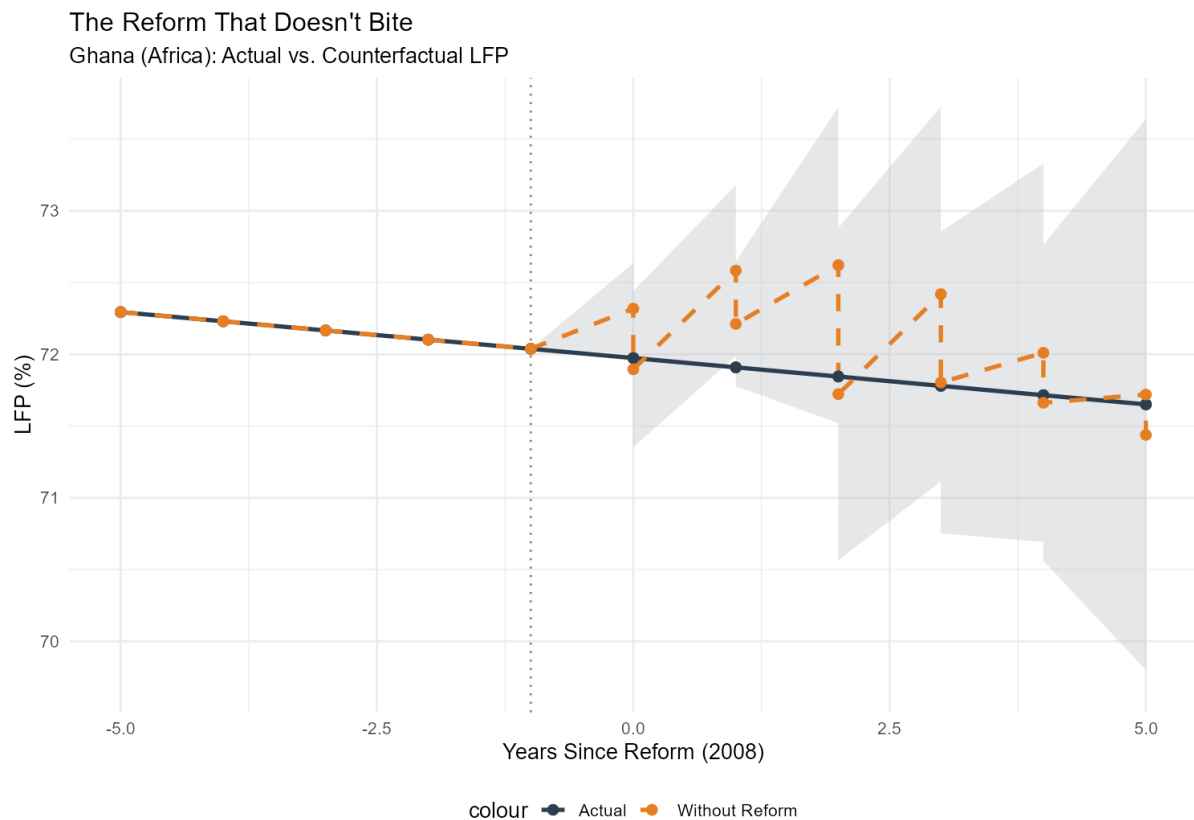
*Notes: Counterfactual LFP constructed using Africa restricted definition coefficients. None of the period-specific effects are statistically significant at  $t = 4$  ( $p = 0.662$ ).*

Ghana's labour force participation rate remained essentially flat throughout the post-reform period, drifting minimally from 72.0 percent in 2007 to 71.7 percent in 2012.

The simulated counterfactual trajectory is virtually identical to the observed path. While the short-run coefficients detect transient contractions at  $k = 0$  and  $k = 1$ , this effect decays completely by  $k = 4$ , shrinking to a statistically insignificant negative 0.30 percentage points. The contrast between Azerbaijan and Ghana is instructive. In formalized labour markets, pension reforms transmit into observable behavioural responses. In high-informality settings, the transmission is weak or absent.

Figure 8 presents the trajectories. The actual and counterfactual lines are nearly indistinguishable, and the confidence bands are wide.

**Figure 8: Ghana Counterfactual**



The contrast between Azerbaijan and Ghana is instructive. In formalized labour markets, pension reforms transmit into observable behavioural responses. In high-informality settings, the transmission is weak or absent. This divergence reflects the structural differences documented throughout this paper: pension coverage, labour market formalization, and enforcement capacity determine whether parametric reforms reach workers and alter their incentives.

### 7.3 Never-Treated Projections

Of the 183 countries in the dataset, 36 recorded no reform activity during the 2000 to 2024 sample period. To translate the panel estimates into forward-looking metrics, this subsection evaluates a hypothetical scenario: what would happen to labour force participation if these static countries activated a major pension reform? A uniform hypothetical reform year of 2010 is imposed, and the restricted definition parameters from the relevant regional sample are applied. Table 7.5 presents projections for four never-treated countries: Chile and Canada (non-Africa), and Botswana and Algeria (Africa).

**Table 7.5: Never-Treated Projections — LFP at t = 4 Under Hypothetical 2010 Reform**

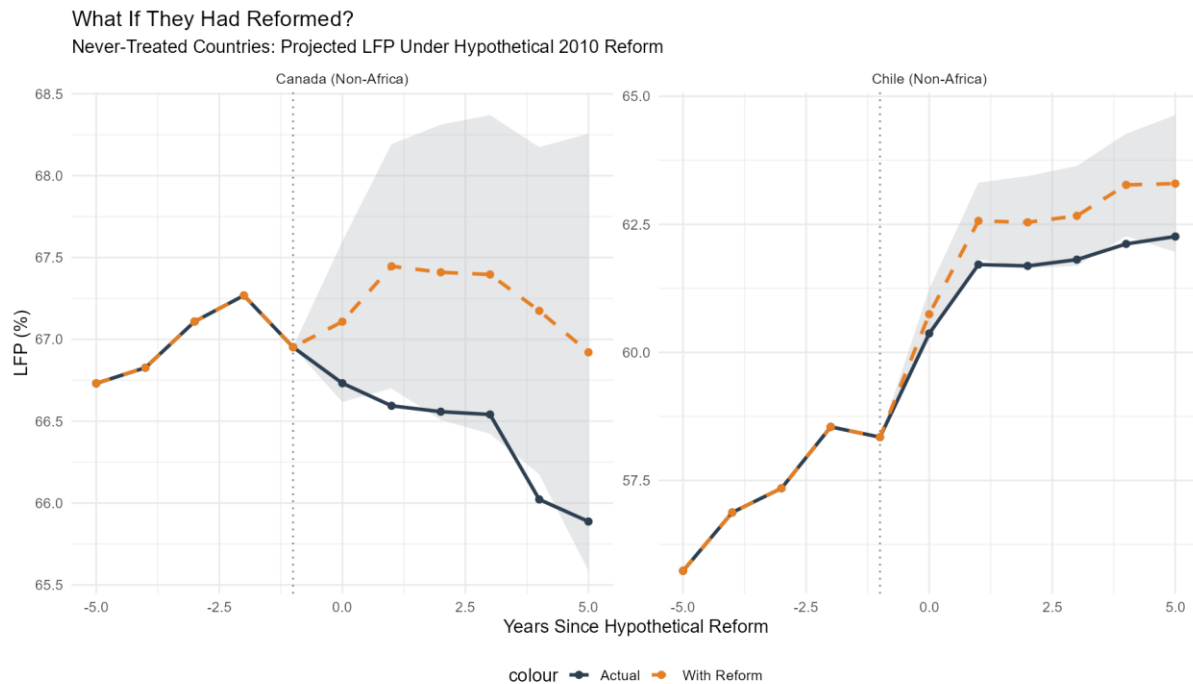
Country	Region	LFP Pre (2009)	Actual LFP (2014)	Projected LFP Under Reform (2014)	Δ from Actual
Chile	Non-Africa	58.3	62.1	61	-1.1
Canada	Non-Africa	67	66	64.8	-1.2
Botswana	Africa	62.7	61.9	61.6	-0.3
Algeria	Africa	43.2	42.6	42.3	-0.3

*Notes: Projected LFP under reform is constructed by adding the period-specific restricted definition coefficients to the actual LFP trajectory. See Section 4.7 for methodology.*

In Chile and Canada, the projected LFP under a hypothetical reform is approximately 1.1 to 1.2 percentage points lower than the actual LFP at t = 4. In Botswana and Algeria, the projected impact is negligible at approximately 0.3 percentage points, reflecting the small and imprecisely estimated Africa treatment effects. These projections are illustrative, not predictive. They assume homogeneous treatment effects within regions, an assumption that cannot be tested for countries that did not actually reform.

Figure 9 presents the trajectories for Chile and Canada with confidence bands.

**Figure 9: Never-Treated Projections**



## 7.4 Implications for Reform Sequencing

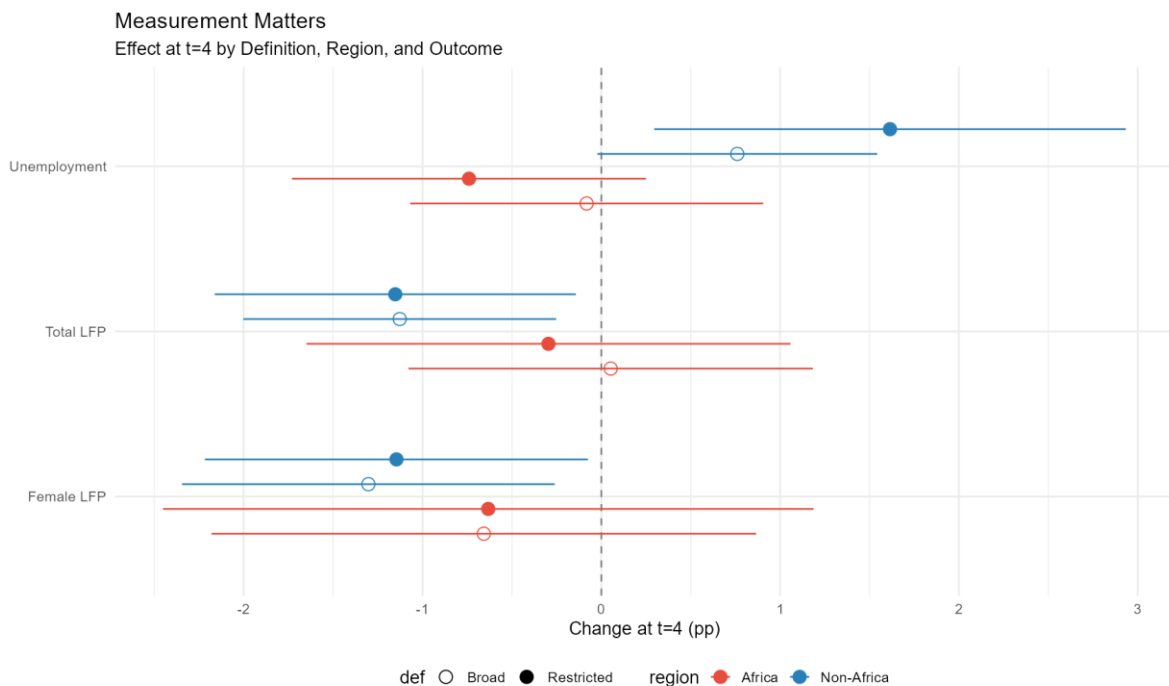
The counterfactual exercises provide insights for the sequencing of pension reform in developing economies. Ghana's 2008 reform serves as an instructive case: despite comprehensive statutory change, the reform failed to generate any long-term behavioural adjustment in aggregate labour supply. Macroeconomic stabilization creates the conditions for reform, but even if reform occurs, the labour supply response remains muted without parallel efforts to expand formal coverage and strengthen enforcement. The policy implication is not that reform is futile. It is that reform alone is insufficient. The sequence matters: coverage expansion and institutional strengthening should accompany or precede parametric reform, not follow it.

## 7.5 Measurement Structure

Figure 10 summarizes the estimated causal parameters at the four-year post-reform horizon across every permutation of definition, region, and labour market outcome. Three patterns are visible. First, the restricted definition consistently yields larger point estimates than the broad definition for participation outcomes in non-African countries. Second, African estimates are consistently closer to zero and bounded by wide confidence intervals, regardless of definition or outcome. Third, the sign and magnitude of the unemployment response depends on region: positive and large in non-Africa, negative or zero in Africa. The measurement matrix confirms the paper's

central argument: before asking what pension reforms do, researchers must ask what is meant by a reform and in what context it is implemented.

**Figure 10: Measurement Matters]**



## 8. Conclusion

This paper has examined the labour market effects of pension reform using a staggered difference-in-differences event study design across a global panel of 183 countries over the period 2000 to 2024. The central contribution is to show that estimated reform effects depend critically on how reform is defined and on the institutional context in which reform is implemented. Three findings emerge.

First, measurement matters. Both the broad definition, which treats any recorded reform activity as a treatment event, and the restricted definition, which isolates major structural and parametric reforms, produce negative effects on labour force participation in non-African countries. But the restricted definition yields larger and more precisely estimated effects, with a cumulative decline of 4.1 percentage points within four years compared to 3.8 under the broad definition. The broad definition pools major reforms with minor administrative amendments, attenuating the estimated response. The choice of definition therefore determines the magnitude and precision of the estimated effect and can influence the policy conclusions drawn from the analysis.

Second, institutional context mediates the response. In African countries, where pension coverage is low and informality is high, the estimated effects are small, short lived, and statistically indistinguishable from zero under both definitions. The

transmission from pension policy to labour supply requires formal coverage, effective enforcement, and a labour market in which statutory parameters bind. Where these conditions are absent, parametric reforms cannot be expected to produce measurable aggregate labour supply responses. Pension reforms are not universally portable policy tools. Their effectiveness depends fundamentally on institutional reach and labour market formalization.

Third, the labour market adjustment is distributionally uneven. In non-African countries, declining participation is accompanied by rising unemployment, consistent with adjustment pressures operating through both demand and supply margins. Female labour supply declines more than total participation in every sample and under every definition. Women bear a disproportionate share of the adjustment, a distributional consequence masked in aggregate analyses.

These findings carry three policy implications. First, researchers and policymakers should distinguish between major reforms that alter incentives and minor amendments that do not. Pooling these distinct forms of policy activity attenuates estimated effects and can lead to misleading conclusions. Second, parametric reform alone is unlikely to shift labour supply in high-informality settings. Expanding coverage, strengthening enforcement, and designing contribution systems suitable for informal workers should accompany or precede parametric reform. Third, the distributional consequences of reform, particularly the gendered pattern of labour supply adjustment, warrant explicit consideration in reform design.

The analysis has limitations that suggest avenues for future research. The event study design identifies average effects across heterogeneous reforms. Disaggregating by reform type, distinguishing retirement age increases from contribution rate changes or benefit formula adjustments, would further clarify the mechanisms at work. The post-reform window is limited to five years. Longer panels would reveal whether the negative participation effects documented here persist, reverse, or give way to supply side responses over longer horizons.

The central insight is that before asking what pension reforms do, researchers and policymakers must ask what they mean by a reform and whether the institutional conditions for transmission are present. Where pension systems cover a small fraction of the workforce, parametric adjustments cannot shift aggregate labour supply. Where coverage is broad, reforms generate significant behavioural adjustments, but those adjustments may include short-term displacement effects that are distributionally unequal. When reform measurement itself determines the estimated effect, the question of what pension reforms do cannot be answered without first asking what we mean by a reform.

## References

- Aguila, E. (2011). Personal retirement accounts and saving. *American Economic Journal: Economic Policy*, 3(4), 1–24.
- Artmann, E., Fuchs-Schündeln, N., & Giupponi, G. (2023). Forward-looking labor supply responses to changes in pension wealth: Evidence from Germany (IZA Discussion Paper No. 16450). Institute of Labor Economics.
- Beccaria, L., & Groisman, F. (2015). Informality and labour market segmentation: The case of Argentina. *International Labour Review*, 154(2), 227–243.
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1), 249–275.
- Cribb, J., Emmerson, C., & Tetlow, G. (2013). Incentives, shocks or signals: Labour supply responses to increases in the female state pension age in the UK (IFS Working Paper No. W13/03). Institute for Fiscal Studies.
- Cribb, J., Emmerson, C., & Tetlow, G. (2014). Labour supply responses to an increase in the legal retirement age for women (IFS Working Paper No. W14/19). Institute for Fiscal Studies.
- De Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- Ebbinghaus, B. (2009). Can path dependence explain institutional change? Two approaches applied to welfare state reform. In *The evolution of path dependence* (pp. 191–218).
- Escobar Latapí, A., & González de la Rocha, M. (2008). Girls, mothers and poverty mitigation in Mexico: Evaluating Progres-Oportunidades. *Journal of Latin American Studies*, 40(1), 67–107.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Gruber, J., & Wise, D. A. (Eds.). (1999). *Social Security and Retirement Around the World*. University of Chicago Press.
- Gruber, J., & Wise, D. A. (Eds.). (2004). *Social Security Programs and Retirement Around the World: Micro-Estimation*. University of Chicago Press.
- Hart, K. (1973). Informal income opportunities and urban employment in Ghana. *Journal of Modern African Studies*, 11(1), 61–89.
- Hernæs, E., Markussen, S., Piggott, J., & Røed, K. (2015). Pension reform and labor supply: Flexibility vs. prescription. *Journal of Public Economics*, 130, 14–25.

- Hinrichs, K., & Jessoula, M. (2012). Labour market flexibility and pension reforms: What prospects for security in old age? In K. Hinrichs & M. Jessoula (Eds.), *Labour market flexibility and pension reforms: Flexible today, secure tomorrow?* (pp. 1–25). Palgrave Macmillan.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7), 1159–1178.
- Pierson, P. (2004). *Politics in Time: History, Institutions, and Social Analysis*. Princeton University Press.
- Roth, J. (2022). Pretest thresholds and confidence intervals after pretesting for parallel trends. *American Economic Review: Insights*, 4(2), 217–232.
- Schmidheiny, K., & Siegloch, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*, 38(5), 695–713.
- Staubli, S., & Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of Public Economics*, 108, 17–32.
- Stewart, F., & Yermo, J. (2009). Pensions in Africa (OECD Working Papers on Insurance and Private Pensions, No. 30). OECD Publishing.
- Stock, J. H., & Wise, D. A. (1990). Pensions, the option value of work, and retirement. *Econometrica*, 58(5), 1151–1180.
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.