

Supply Without Demand: Japan's Pension Reform and Elderly Male Labor Force Participation

Terry S. Lo

Cardiff Sixth Form College, Trinity Court, 21-27 Newport Rd, Cardiff CF24 0AA, Wales, UK of Great Britain and Northern Ireland;
terrlo20090326@gmail.com

ABSTRACT: This paper examines the effect of Japan's 2001 pension eligibility age increase and 2006 Elderly Employment Stabilization Law on male labor force participation for ages 60–64. Using synthetic control and synthetic difference-in-differences with a 27 and 19-country panel dataset spanning 1991–2022, this paper finds no detectable effect of the 2001 pension reform alone, but suggests a positive effect for the 2006 Elderly Employment Stabilization Law. The ATT flips from negative to positive, a swing of nearly 11 percentage points. When using the SDID estimator, the effects of both policies were not significant.

KEYWORDS: Economics, Public Economics, Pension Reform, Japan, Labor Force Participation

■ Introduction

Japan, with a third of its population over 65, has the highest share of elderly in the world.¹ Moreover, more than 1 in 10 people in the country are now aged 80 or older.¹ This shift in demography is caused by the following three key factors: the post-WWII baby boom between 1947 and 1949, when births exceeded 2.5 million for three consecutive years;² a declining fertility rate, which fell to a record low of 1.15 in 2024,² well below the replacement rate of 2.1;³ and a rising life expectancy of their citizens, with Japanese women living an average of 87.13 years, which is the highest in the world and with men living 81.09 years.⁴ This rapid aging process piles pressure on Japan's pension system, as a shrinking working-age population has to support a growing number of retirees' income.

Japan's aging population means that there will be fewer workers supporting the pension beneficiaries, as the ratio of working-age people to retirees has experienced a constant decline from 9.5 working-age people per retiree in 1960 to 2.6 by 2009.⁵ This placed a growing financial burden on Japan's pay-as-you-go pension system. Therefore, the government enacted pension reform to the eligibility age, which aimed to reduce the total period of benefit payouts, thus mitigating the fiscal pressure on the system.

Japan's pension reform was legislated in 1994 and 2000, increasing the eligibility age from 60 to 65 for flat-rate and wage-proportional pensions, respectively. Both reforms took a gradual approach. The flat-rate component reform was implemented in 2001 for men, with the eligible age increasing one year every three years. However, this created a gap between the new age for the flat-rate component of the pension and the mandatory retirement age, which remained at 60 in most firms, leaving workers without employment and with reduced pension income during those years. To address this, the Elderly Employment Stabilization Law was revised in 2006, requiring firms to retain older workers until the new eligibility age.⁶ Unless otherwise specified, 'pension reform' in this paper refers to

the 1994 reform that increased the flat-rate component eligibility age, implemented in 2001-5

This study hypothesizes that raising the pension eligibility age increased male labor force participation for ages 60–64. The delay of eligibility reduces lifetime pension wealth and makes leisure relatively more expensive, both of which encourage continued work. The theoretical framework section develops this logic further.

Previous studies examined Japan's pension reform using within-country comparisons.^{7,8} These studies found that the reform had a positive impact on employment rates, especially after the 2006 law addressed the gap between retirement age and pension eligibility, where it increased employment rates for men aged 60–61 by 2 to 3 percentage points,⁷ and that combining the eligibility age increase with demand-side enforcement raised employment at age 63 by approximately 4 percentage points.⁷ However, within-country comparisons may be subject to spillover effects. For example, the same eligibility age increase applied to women five years later, meaning that men in supposedly untreated cohorts may have adjusted their labor supply in response to the reduced household pension income from their spouse's delayed eligibility. Cross-country methods avoid this possible contamination. To date, no study has used cross-country comparisons to investigate the effect of the reform. This study addresses this absence of cross-country evidence by using synthetic control and synthetic difference-in-differences to examine the effect of Japan's pension reform on labor force participation. These methods construct a counterfactual using other countries, avoiding within-country contamination.

The pension reform and the 2006 Elderly Employment Stabilization Law affected the labor force participation rate in different ways. The pension reform operates on the supply side by reducing pension wealth and increasing the incentive to work. However, the majority of firms in Japan had their mandatory retirement age set as 60, which created a demand-side constraint, as employers could force workers out regardless of their willingness to continue working.

To disentangle supply-side from demand-side effects, this paper additionally estimated a 2006 treatment year corresponding to the Revised Elderly Employment Stabilization Law. The key pattern is that the treatment effect flips from negative under the 2001 treatment to positive under the 2006 treatment, a swing of nearly 11 percentage points. This suggests that pension reform required complementary employment legislation to produce labor force participation gains. This finding has broader policy relevance, as many developed countries are currently considering or implementing similar increases to pension eligibility ages.

Institutional Background:

The system:

Japan's public pension system operates on a two-tier structure. The first tier is the National Pension Insurance (NPI, Kokumin Nenkin), which provides a flat-rate basic benefit to all residents regardless of lifetime earnings. The eligibility age for NPI is 65 years old. The benefit from NPI is called the Old-Age Basic Pension. Everyone receives the Old-Age Basic Pension from age 65.

The second tier consists of earnings-related schemes for employees. The Employees' Pension Insurance (EPI, Kosei Nenkin) provides earnings-related benefits to employees; the eligibility age for EPI was 60 years old pre-reform. The Mutual Aid Insurance (MAI) covers government employees and private school teachers. The MAI followed the same eligibility age reform schedule as EPI and was integrated into EPI in October 2015.⁵ At 65 years old, 45.5% of the population receive NPI only, 48% receive NPI plus EPI (wage-proportional component), and 6.5% receive the Old-Age Basic Pension plus MAI wage-proportional component.⁹ This study focuses on EPI as it covers the largest share of employees. Within the EPI, benefits are divided into two components:

1. Flat-rate component - A fixed amount of money equivalent to the NPI basic pension. This component was designed to substitute for the NPI for retirees under the age of 65. Once they reach 65, the EPI flat-rate component stops and is replaced by the NPI.

2. Wage-proportional component - based on career average earnings. The benefit from EPI is called the Old-Age Employees' Pension. Before age 65, EPI members receive the Special Old-Age Employees' Pension (flat-rate + wage-proportional). From age 65, they receive the Old-Age Basic Pension (from NPI) plus the Old-Age Employees' Pension (wage-proportional component only from EPI).

Before the reforms, EPI beneficiaries could claim both components starting at age 60 through a transitional arrangement called the "Special Payment of the Employees' Old-Age Pension" (TOKUBETSU SHIKYUNO ROREI KOSEI NENKIN).

The reform package (1994 and 2000):

Japan's pension eligibility reform was implemented through two major legislation acts, creating a phased transition that separates the flat-rate component from the wage-proportional component.

1994 Reform Act:

The 1994 Pension Reform Act legislated the increase in eligibility age for the flat-rate component of the EPI from age 60 to 65. The implementation began in April 2001 for male workers, with the eligibility age increasing by one year every three years, as shown in Table 1:⁵

Table 1: Implementation schedule for flat-rate component eligibility age increase (1994 reform, male workers).

Birth Cohort	Eligibility Age	First Affected Year
Born before April 2, 1941	60	-
April 2, 1941 - April 1, 1943	61	2001
April 2, 1943 - April 1, 1945	62	2004
April 2, 1945 - April 1, 1947	63	2007
April 2, 1947 - April 1, 1949	64	2010
Born after April 1, 1949	65	2013

Note: This reform applied to female workers as well, with a 5-year time lag. The flat-rate component eligibility age for females increased with the same pattern starting in 2006.⁵ The flat-rate eligibility age rises in one-year steps from 60 for cohorts born before April 1941 to 65 for those born after April 1949.

The 2000 Reform Act:

The 2000 Pension Reform Act (passed in March 2000) extended this approach to the wage-proportional component. Implementation began in April 2013 for male workers (and in 2018 for female workers), with the eligibility age increasing by one year every three years from 60 to 65, following the same phased pattern as the flat-rate component (completing in 2025 for men).⁵

The Earnings Test (Zaishoku Rorei Nenkin):

Japan's pension system includes an earnings test (Zaishoku Rorei Nenkin) that reduces pension benefits for workers aged 60–64 whose combined monthly earnings and pension exceed 280,000 yen,¹⁰ with an effective marginal tax rate of 50%.⁸ For high earners, pensions are fully suspended. This mechanism interacts with pension reform by reducing the incentive to work while receiving pensions, but it is not the focus of this study.

The 2006 Elderly Employment Stabilization Law:

The Gap Problem. The pension reforms created a growing gap between the flat pension eligibility age and the mandatory retirement age, which remained at 60 in most firms. This gap emerged as a serious social concern,⁷ as workers faced a period with neither employment nor sufficient pension income.

The 2006 Legislative Response:

To address this problem, the government revised the Elderly Employment Stabilization Law in 2004, with implementation beginning in April 2006. Beginning with workers who turned 60 in April 2006 (those born in April 1946), employers were legally required to take at least one of the following measures:⁶

1. Raise the mandatory retirement age to the pension eligibility age.
2. Abolish mandatory retirement entirely.
3. Introduce a continued employment system.

Firm Response and Employment Effects. In practice, more than 80% of firms chose the third option,⁶ where they offer reemployment contracts to workers after mandatory retirement. These contracts last until the workers reach the pension eligibility age. For example, it is age 63 for those born in 1946, age 65 for those born after April 1949. Reemployed workers usually receive lower wages. Kondo and Shigeoka found that this legal enforcement increased the employment rate for men in their early 60s.⁷

Scenarios. To better understand how the reforms affected different workers, four scenarios are presented below. The experience of each cohort depends on their birth year, pension eligibility age, the gap, and the 2006 law.

The first scenario is NPI only. This applies to workers who are not employed in the private or public sector, such as the self-employed. They are not affected by the EPI eligibility age reforms and receive the Old-Age Basic Pension from NPI starting at age 65. Since they are not subject to mandatory retirement, they do not face the gap problem.

The second scenario is EPI with reform impact but no law addressing the gap. This represents the cohort born between April 2, 1941, and April 1, 1946. Here we use a male born in April 1943 as an example. He turned 60 in 2003; the flat-rate eligibility age was 62, meaning that he could only receive his wage-proportional component. He had to wait until 2005, when he reached 62, to start receiving his flat-rate component. The 2006 law was not yet legislated, so his employer was not required to offer reemployment. This cohort faced the gap: forced out of employment at 60 with reduced pension income for two years.

The third scenario is EPI with reform impact and gap addressed by the 2006 law. This represents the cohort born between April 2, 1946, and April 1, 1949. Here we use a male born in April 1946 as an example. He turned 60 in April 2006 and would have faced mandatory retirement. His flat-rate eligibility age was 63, so at 60, he could only receive the wage-proportional component. The 2006 revision of the Elderly Employment Stabilization Law was now in effect, requiring his employer to offer reemployment until his flat-rate eligibility age, which was 63 in his case. He was therefore rehired at lower wages. At 63, he received both components and ended his reemployment contract. At 65, the flat-rate component was replaced by the Old-Age Basic Pension.

The fourth scenario is EPI with both reforms (flat-rate and wage-proportional delay). This represents the cohort born after April 1, 1949. Here we use a male born on April 15, 1953, as an example. He turned 60 in 2013. The flat-rate eligibility age was 65, and the wage-proportional eligibility age was 61. The 2006 law applied, and he was offered a reemployment contract that lasted until age 65. In 2014, he turned 61 and started receiving the wage-proportional component while continuing to work. In 2018, he turned 65 and retired. He then received the Old-Age Basic Pension (NPI) plus the Old-Age Employees' Pension (wage-proportional from EPI).

Across these four scenarios, the critical determinant of each cohort's experience is the gap between mandatory retirement at 60 and the rising pension eligibility age, and whether the

2006 law was in effect to bridge that gap through employer-mandated reemployment.

Theoretical Framework:

Implicit Tax on Continued Work. Early retirement age (ERA) is the age when Social Security benefits are first available. Social Security Wealth (SSW) is the present discounted value of all future pension benefits a worker is entitled to receive based on their work history. When a worker considers working one more year instead of retiring, their SSW will change. This change is called the accrual.¹¹

The accrual can be either positive or negative, depending on the country's pension rules. If a country's pension system offers no reward for delayed claiming of pension, working an additional year will imply a person has to give up one year of pension payments without gaining higher benefits in the future. This creates a negative accrual as SSW is reduced. A negative accrual acts like a tax on working because it reduces the value of an individual's salary, which means part of what they earn will be offset by the pension income that they have given up. For example, a worker at age 60 who would receive 120 annually in pension benefits decides instead to work and earn 360. Of this salary, 120 is effectively offset by the foregone pension, which becomes an implicit tax on continued work. In this example, the implicit tax rate is $120/360$, 33% of annual earnings.

When the accrual is expressed as a percentage of annual earnings, it becomes the implicit tax rate. Tax force is the sum of implicit taxes from the early retirement age to age 69.¹¹ A country with a high tax force penalizes continued work after the early retirement age, which means workers who delay retirement will experience a loss in pension wealth. In contrast, workers in countries with low tax force face little penalty for working longer; therefore, they have more incentive to stay in the workforce.

Substitution and Income Effects. When the minimum age for pension eligibility in Japan increases, two effects operate on labor supply: the substitution effect and the income effect.¹² To illustrate, consider a 60-year-old worker deciding whether to work or retire. Before the reform, if he retired, he received a full pension of 100. If he works, he earns 150 in salary and receives 100 in pension, but the earnings test reduces his pension by 50, giving him 200 total (150 salary + 100 pension – 50 reduction). The gain from working is 100. After the reform, pension eligibility is delayed. If he retires, he receives nothing. If he works, he earns 150 with no pension reduction. The gain from working is 150. The reform increased the return to working from 100 to 150. Leisure (not working) became relatively more expensive (in terms of opportunity cost); therefore, it encourages work. This is the substitution effect.

The income effect operates through lifetime pension wealth. As the eligibility age increases, the total pension wealth a person can receive over their lifetime is reduced, since they receive the same annual amount but starting later. This reduction acts as a negative income shock. Since leisure is a normal good, a decrease in wealth leads individuals to consume less leisure, meaning they work more.

Japan's Position in the International Context. The NBER International Social Security project examined pension incentives across 12 developed countries.¹¹ A key finding is the substantial variation in implicit tax rates and how it affects the incentives for working after ERA. As Coile *et al.* state:¹¹ "Working from the ERA to age 69 results in a loss of SSW equivalent to about 1.6 years of earnings in the US and Japan, the two countries whose systems are the closest to neutral with respect to work. By contrast, in Belgium, France, Italy, and the Netherlands, working from the ERA to age 69 results in a loss of SSW equivalent to 7.2–9.2 years of earnings."¹¹

Japan had one of the lowest tax force values among developed countries, with 160%, compared to 800–900% in France, Belgium, and the Netherlands.¹¹ At 900%, nearly all of a worker's earnings are offset by lost pension wealth, meaning the financial gain from continued work is close to zero. 160% means that the cumulative pension wealth a typical worker forgoes by working from the ERA to age 69 is equivalent to 1.6 years of his earnings. This means Japanese workers faced relatively little penalty for working past the ERA. This low penalty for working longer can be reflected in Japan's relatively high labor force participation rate; approximately 71–76% of men aged 60–64 were in the labor force during the 1990s, compared to less than 20% in France and Belgium.¹¹

The Health Capacity Ceiling. Health capacity is one of the limiting factors for the elderly to continue to stay in the workforce. Research on the health capacity to work provides additional context. Wise estimated that approximately 80% of men aged 60–64 have sufficient health to work,¹¹ by combining the estimated effect of health on employment for younger individuals with the actual health of older individuals in order to predict the capacity to work. As Coile *et al.* report:¹¹ "the estimated capacity to work (total height of the bar) of men is fairly similar across countries, at roughly 80 percent, while the share actually working (shaded portion) ranges from about 15 to 80 percent."¹¹ It is worth noting that the countries the quote above is referring to are 12 Western developed countries.

Japan's pre-reform 60–64 cohort's LFPR of 71–76% was relatively high compared to other developed countries, though still several percentage points below this estimated health capacity ceiling.

Implications for Expected Treatment Effect. These considerations suggest that Japan's pension reform may produce a smaller effect than similar reforms elsewhere. With implicit tax rates already low and labor force participation already near the health capacity ceiling, there was limited room for improvement. This does not mean the reform had no effect; rather, it suggests that detecting an effect requires careful consideration of the counterfactual.

Furthermore, supply-side incentives only translate to higher employment if workers can find jobs. Japan's mandatory retirement system was set at age 60 before the 2006 Law, creating a demand-side constraint where even workers who wanted to continue working would be forced to retire. The 2006 Elderly Employment Stabilization Law addressed this by requiring firms to retain workers until the pension eligibility age. This

suggests that the pension reform's effect may depend on whether demand-side barriers are addressed alongside.

Many developed countries also reformed their pension systems during this period. If other developed countries were simultaneously reducing their implicit taxes, Japan's LFPR may rise but show no differential effect relative to other reforming nations.

An additional contextual factor is Japan's prolonged economic stagnation during the 1990s. The pre-treatment period (1991–2000) coincides with the Lost Decade, during which GDP growth stagnated, unemployment rose from approximately 2% to 5%,¹³ and deflation set in. These macroeconomic conditions likely depressed labor force participation during the pre-treatment period and may affect the counterfactual trajectory estimated by SDID.

Literature review:

Shimizutani and Oshio examined the earnings test for workers aged 65–69 by analyzing the elimination of the earnings test in 1985 and its reinstatement in 2002.⁸ They employed the decomposition method proposed by DiNardo, Fortin, and Lemieux¹⁴ with bunching analysis. They found that its elimination in 1985 increased labor supply for the elderly, especially for males, with little evidence that reinstating the test in 2002 affected the labor supply of the elderly. Kondo and Shigeoka focused on workers aged 60–64.⁷ They employed a difference-in-differences approach with varying treatment and control cohorts to measure the impacts of the policies. They found that the 2006 Employment Law increased employment rates, and that the effect of raising the pension eligibility age was larger when combined with demand-side enforcement. Okumura and Usui examined how pension reform affected expectations about future benefits,⁹ using survey and regression models. They found that individuals responded to the reform by increasing savings as it reduced expectations of future benefits. This suggests that pension reform affects behavior even before workers reach retirement age.

All of these studies used within-country comparisons – comparing treated units with other cohorts in the country. However, within-country quasi-experimental approaches require strong assumptions that policy affecting one group does not spill over to other groups.¹¹ For example, as the reform took place, workers expecting future pension changes may adjust their employment decisions in advance, potentially contaminating the control group, which might distort the result.

To date, no study has used cross-country comparisons to examine Japan's pension reform. Cross-country methods construct a counterfactual using other countries by giving them weight using data-driven optimization methods, avoiding within-country contamination.

The International Social Security project provides important context for interpreting pension reform effects across countries. They discovered that reforms across developed countries since the 1980s have generally reduced implicit taxes on continued work,¹¹ which led to rising employment at older ages. It also shows that these reforms explain a substantial share of the increase in older worker employment. Crucially, this was a global

phenomenon among developed nations, not unique to Japan. This suggests that any cross-country analysis of Japan's reform must account for the fact that control countries may also have been reforming their pension systems during the same period.

■ Methods

To implement synthetic control and synthetic difference-in-differences, a panel dataset is constructed that spans from 1991 to 2022. The panel dataset combines data from various sources. Male labor force participation rates (LFPR) for ages 60–64, the outcome variable, are obtained from ILOSTAT, the International Labor Organization's statistics database.¹⁵ Covariates, which include GDP per capita, unemployment rate, and population aged 65 and above, are drawn from the World Bank World Development Indicators.¹³ ILOSTAT provides labor statistics across 189 countries. WDI covers 217 economies.

The original datasets contained over 100 countries. However, many were dropped during data cleaning due to missing values in the outcome variable or covariates, leaving 27 countries with complete data for the study period. The final sample includes 27 countries: Australia, Austria, Belgium, Canada, Chile, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, Turkey, and the United States. However, not all countries are suitable to be in the donor pool. Two synthetic control specifications were run, one with 27 and one with 19 countries. Eight countries are excluded from the main specification to ensure the donor pool is comparable to Japan. Chile, Mexico, and Singapore are excluded because their pension systems are structurally different from Japan's pay-as-you-go system: Chile and Mexico operate privatized individual account systems,¹⁶ and Singapore uses a provident fund.¹⁷ Korea is excluded because its pension system was only established in 1988 and did not reach universal coverage until 1999,¹⁸ leaving insufficient pre-treatment data. Turkey is excluded due to high labor market informality. Iceland and Luxembourg are excluded due to small populations (all under 1 million)¹³ and distinct labor market characteristics, such as Iceland's near-full employment. Norway is excluded due to its oil wealth. This leaves 19 developed countries as the main donor pool. Unlike Abadie *et al.*, who used SC, which used a subset of pre-treatment outcome years, all pre-treatment outcome years were used as predictors for synthetic control to maximize pre-treatment fit.¹⁹

This paper focuses on males because the pension reform for females began five years later, leaving less post-treatment data available. The 60–64 age group is selected because this cohort was directly affected by the increase in pension eligibility age from 60 to 65. The treatment year is 2001, when the eligibility age for the flat-rate component first increased from 60 to 61.

Although the reform was implemented gradually over 12 years, 2001 is used as the treatment year because it marks when the policy first began to bind and is the first cohort to face delayed eligibility. Anticipation effects are unlikely to threaten pre-treatment trends because the cohorts aged 60–64 during

1991–2000 were unaffected by the reform. Workers aged 55–59 during 1994–2000 who anticipated the reform may have adjusted their savings or employment plans in advance. However, such anticipatory increases in labor supply would, if anything, decrease post-treatment LFPR, biasing against finding a positive treatment effect. Since the reform is phased in over 12 years, the ATT averages across different treatment intensities, which may underestimate the full reform effect if the effect increases with the size of the eligibility age increase. This approach follows the paper that proposed synthetic control,¹⁹ which similarly treated California's phased Proposition 99 tobacco control program as a single 1988 treatment.

The synthetic control method constructs a counterfactual and compares it with the treated unit.¹⁹ The difference between them is the effect of the reform. The counterfactual is constructed by assigning weights to donor countries to match Japan's pre-treatment LFPR level. Weights are chosen to minimize the difference between Japan and the synthetic control in pre-treatment periods, using both pre-treatment outcome values and covariates (GDP per capita, unemployment rate, and population aged 65+). A combination of countries often provides a better comparison for the region exposed to the intervention than any single country alone. However, when restricted to comparable developed countries, synthetic control fails to achieve adequate pre-treatment fit, where only Switzerland remains in the donor pool with 100% weight; other developed countries, including New Zealand and Denmark, remained in the donor pool but received zero weight. Indicating that Japan is an outlier outside the convex hull of donor countries. Synthetic difference-in-differences is therefore used for further analysis.²⁰ Unlike synthetic control, SDID uses only the outcome variable and matches on trends rather than levels. The SDID method reallocates weight and attempts to mimic pre-exposure trends using a combination of countries; doing so can weaken the reliance on the parallel trend assumption that traditional DID relies on. The method then assumes that Japan's trend would have been the same as the counterfactual's trend. Standard errors are computed using the placebo variance estimator proposed by Arkhangelsky, Athey, Hirshberg, Imbens, and Wager (2021), which is designed for settings with a single treated unit.

Both methods rely on several assumptions. First, the assumption of no interference between units requires that Japan's treatment does not affect outcomes in donor countries and that no spillover occurs between treated and control units.¹⁹ Second, no similar treatment should be implemented in the donor pool during the study period. For synthetic control, it requires that the treated unit's pre-treatment values can be matched by a weighted combination of donors, which means Japan has to lie within the convex hull of the donor pool. When this condition is violated, which is the case in this study, "we would not recommend using a synthetic control" (p. 495).¹⁹

Both methods assume no similar treatment occurred in the donor pool. However, many developed countries also reformed their pension systems during the period of study, which affects the interpretation of the estimates.

■ Results and Discussion

The increase in eligibility age:

This section presents the results of the synthetic control and synthetic difference-in-differences. Two synthetic controls were run with progressively restricted donor pools. Model 1 includes all 27 countries with complete data. Model 2 excludes Chile, Mexico, Singapore, Korea, Turkey, Iceland, Norway, and Luxembourg, leaving 19 developed countries. Table 2 reports the donor weight for each run, and Table 3 reports the pre-treatment fit.

Table 2: Synthetic control donor weights by model specification.

Country	Model 1 (27)	Model 2 (19)
Mexico	0.429	N/A
Turkey	0.246	N/A
New Zealand	0.089	0
Iceland	0.129	N/A
Chile	0.067	N/A
Norway	0	N/A
Denmark	0.04	0
Switzerland	0	1
All others	0	0

Note: N/A indicates the country was excluded from the donor pool. Weights sum up to 1 within each model. Only countries with non-zero weights in at least one specification are shown. Weight concentrates on emerging-market donors in Model 1 (Mexico, Turkey, and Chile together hold 74.2%), whereas in Model 2, all weight is placed on Switzerland, showing that no combination of developed-country donors can reproduce Japan.

Table 3: Synthetic control pre-treatment fit.

Model	Donor Pool	RMSPE
1	27 countries	0.544
2	19 countries	6.156

Note: RMSPE = Root Mean Squared Prediction Error for pre-treatment period (1991–2000). Lower values indicate a better fit. Pre-treatment RMSPE rises from 0.54 in Model 1 to 6.16 in Model 2, confirming that the restricted developed-country donor pool fits Japan poorly.

Table 2 shows that as the donor pool becomes more restricted, the weights concentrate on fewer countries. In Model 1, weights concentrate on 6 countries: Mexico 42.9%, Turkey 24.6%, Iceland 12.9%, New Zealand 8.9%, Chile 6.7%, and Denmark 4.0%. Notably, emerging markets (Mexico, Turkey, Chile) account for 74.2% of the total weight. In Model 2, the weight is concentrated on one country, Switzerland, with 100%, indicating that Japan cannot be matched by a weighted combination of the remaining developed donor countries. Table 3 further confirms this; it shows the RMSPE increases from 0.54 in Model 1 to 6.16 in Model 2, indicating substantially poorer pre-treatment fit. The good fit in Model 1 (RMSPE = 0.54) is achieved only by relying heavily on emerging markets, which are fundamentally different from Japan in economic structure. While pre-treatment outcome values show excellent fit, structural covariates show substantial imbalance: Japan's GDP per capita is 67% higher than that of synthetic Japan (Model 1), and Japan's elderly population share is more than double that of synthetic Japan. This suggests that Japan lies outside the convex hull of developed country donors; its pre-treatment LFPR level is too distinct to be replicated by any combination of com-

parable countries. Synthetic control is not recommended when this condition is violated.¹⁹

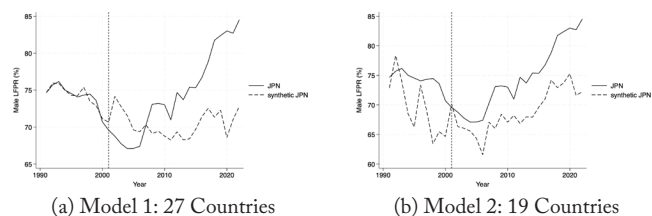


Figure 1: Synthetic control results by model specification.

Note: The solid line represents Japan. The dashed line represents synthetic Japan. The vertical dashed line indicates the treatment year (2001). Model 1 (27 countries) closely matches Japan before 2001 but relies heavily on emerging-market donors, while Model 2 (19 developed countries) diverges across the whole period, indicating that Japan lies outside of the convex hull of comparable developed countries.

Figure 1 confirms these patterns visually. Model 1 tracks Japan closely pre-treatment, but the post-2001 divergence must be interpreted with caution, given the emerging market composition. Model 2 shows a poor fit, diverging throughout the entire period, confirming that the convex hull condition is violated. Given the convex hull violation, SDID is used as the main estimator.

Table 4: Synthetic difference-in-differences results.

Specification	ATT	SE	p-value	95%CI
19 countries	-8.34	6.65	0.210	[-21.37, 4.70]
27 countries	-5.36	7.79	0.522	[-20.64, 9.91]

Note: ATT = Average Treatment Effect on the treated unit. SE = Standard Error. p-value = probability value for the null hypothesis that ATT = 0. 95% CI = 95% Confidence Interval. The countries include Japan. Both specifications yield negative ATTs (-8.34 and -5.36 percentage points) that are not statistically significant, so no effect of the 2001 reform is detected.

Table 4 presents the SDID results. The ATT for the main specification with 19 developed countries is -8.34 percentage points, with a standard error of 6.65, which indicates that, on average, Japan's LFPR was 8.34 percentage points below what the counterfactual would have predicted after the reform. However, the negative point estimate does not imply the reform reduced labor force participation. The standard error of 6.65 shows that there is uncertainty in the ATT estimate. The p-value of 0.210 is greater than the conventional threshold of 0.05, indicating that the effect is not statistically significant. The 95% CI [-21.37, 4.70] includes zero, implying that the true effect cannot be distinguished from no effect. The wide confidence interval includes both negative and positive values, meaning the true effect could range from a 22 percentage point decrease to a 5 percentage point increase. The robustness check using all 27 countries gives similar results (ATT = -5.36, p = 0.522), confirming that the null finding is robust to a change in the donor pool.

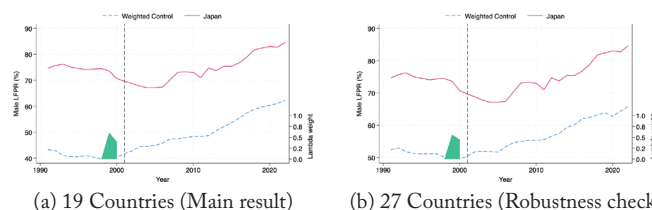


Figure 2: Synthetic difference-in-differences results.

Note: The red solid line is Japan. The blue dashed line is the weighted control. The green shaded area is the time weights. The vertical dashed line indicates the treatment year (2001). In both donor pools, Japan and the weighted control follow similar upward trends after 2001, with Japan slightly weaker, and the estimated effect is not statistically significant.

Figure 2 depicts the SDID results. In both graphs, Japan and control follow similar upward trends in the post-treatment period, with Japan showing a slightly weaker upward momentum. The two graphs show nearly identical patterns, confirming that the null finding is robust to a change in the donor pool.

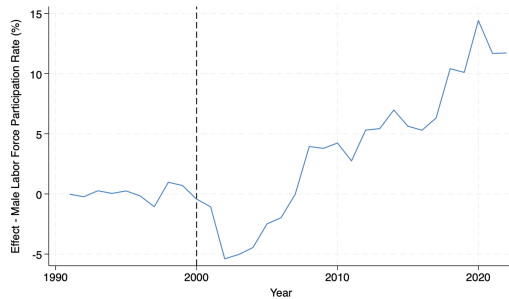


Figure 3: Year-by-year treatment effect (27 countries SC).

Note: Effect is Japan minus synthetic Japan. The vertical dashed line indicates the treatment year (2001). The effect was negative from 2001 to 2006, then turned positive from 2008 onwards, coinciding with the implementation of the 2006 Elderly Employment Stabilization Law.

Year-by-year effects from synthetic control show that from 2001 to 2006, Japan's LFPR was approximately 1.0 to 5.4 percentage points below the synthetic control. Starting in 2008, Japan exceeded synthetic control, reaching approximately 12 percentage points above by 2022. This timing coincides with the 2006 Elderly Employment Stabilization Law, which required firms to retain workers until the pension eligibility age. This suggests that the pension reform alone may have had little effect, but when combined with demand-side enforcement, it produced positive effects. However, individual year effects are not statistically significant ($p > 0.05$ for most years), so this interpretation is suggestive rather than conclusive.

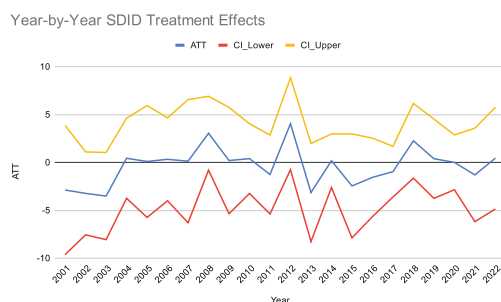


Figure 4: Year-by-year SDID treatment effects (19 countries).

Note: ATT is the year-by-year SDID estimate for the 19-country specification. CI_Lower and CI_Upper are the 95% confidence interval counts. Effects fluctuate around zero with no individual year significant at the 5% level.

Year-by-year SDID estimates show that Japan's LFPR was below the counterfactual in 2001-2003, with effects of -2.88, -3.23, and -3.51 percentage points, respectively, though no individual year was statistically significant at the 5% level. From 2004 onwards, effects fluctuated around zero with no sustained positive or negative trend. The closest to significance was 2012

(ATT = +4.05, $p = 0.099$). No individual year effect was statistically significant, consistent with the overall null finding.

The year-by-year patterns differ between Figure 3 (synthetic control with 27 countries) and Figure 4 (SDID with 19 countries). In Figure 3, Japan exceeded its synthetic control after 2008, while in Figure 4, effects fluctuate around zero throughout. This discrepancy arises from two sources. Firstly, their donor pool is different; the 27-country SC relies heavily on emerging markets (Mexico 42.9%, Turkey 24.6%), while the 19-country SDID excludes these countries. Secondly, the methods they used are different: SC matches on levels, while SDID matches on trends. The SDID results with the restricted donor pool of developed countries are more appropriate for policy interpretation.

The Elderly Employment Stabilization Law:

To test whether demand-side legislation was necessary for the pension reform to take effect, the analysis is re-estimated using 2006 as the treatment year. This corresponds to the implementation of the Revised Elderly Employment Stabilization Law, which required firms to offer rehire contracts until the pension eligibility age.

Synthetic control with 27 countries assigned weight to Turkey (37.1%), Mexico (26.9%), Iceland (21.9%), Chile (9.4%), and Norway (3.8%), with an RMSPE of 1.09. Figure 5 shows the SC path plot. During the pre-treatment period (1991-2005), there was a negative impact on LFPR due to the Lost Decade. After 2006, Japan's LFPR crossed above the synthetic control and diverged progressively, reaching approximately 18 percentage points above by 2020.

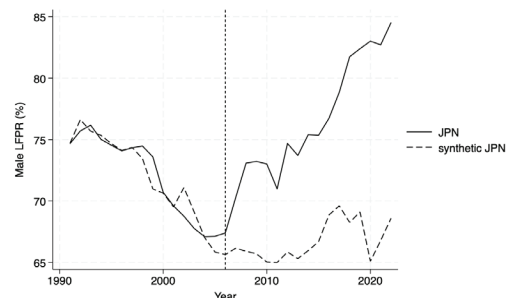


Figure 5: Synthetic control 2006 (27 countries).

Note: The solid line is Japan, and the dashed line is synthetic Japan. The vertical dashed line indicates the treatment year (2006). After 2006, Japan crossed above synthetic control and diverged, reaching approximately 18 percentage points above by 2020.

This sign flip is the most notable pattern in the results. Under the 2001 treatment, the SDID ATT is -8.34 percentage points; under the 2006 treatment, the SDID ATT is +2.50 percentage points ($SE = 7.92$, $p = 0.752$). Although the 2006 estimate is not statistically significant on its own, the reversal from negative to positive, a swing of nearly 11 percentage points, is consistent with the interpretation that pension reform alone was insufficient without complementary employment legislation, where pension reform pushed labor supply without addressing demand-side constraints, while the 2006 law provided the employment mechanism through which increased supply-side incentives could translate into actual participation.

It is important to note that although the sign flip is consistent with the previous result, the size of the coefficients is not significant. This suggests no effect of the policy following the SDID. This leads us to be cautious in interpreting the results from the synthetic control.

Synthetic control placebo-based inference for the 2006 treatment (Table 5) shows effects growing from +4.12 in 2007 to +8.66 percentage points by 2015 (standardized $p = 0.208$) and +15.93 by 2022 (standardized $p = 0.042$), which is statistically significant at the 5% level. The gradual emergence of the effect is consistent with the phased implementation of both reforms interacting over time.

Table 5: Placebo Inference for 2006 synthetic control (27 countries).

Year	Effect (pp)	p-value	Standardised p
2006	+1.77	0.708	0.458
2007	+4.12	0.458	0.208
2008	+7.18	0.25	0.125
2009	+7.52	0.208	0.125
2010	+7.98	0.25	0.083
2011	+5.99	0.458	0.167
2012	+8.83	0.375	0.125
2013	+8.39	0.33	0.125
2014	+9.44	0.375	0.167
2015	+8.66	0.417	0.208
2016	+7.86	0.458	0.167
2017	+9.24	0.333	0.208
2018	+13.49	0.167	0.042
2019	+13.28	0.125	0.125
2020	+17.92	0.083	0.000
2021	+15.92	0.083	0.000
2022	+15.93	0.083	0.042

Note: pp refers to percentage points. Standardized p-values divide each unit's post-treatment effect by its pre-treatment RMSPE, following Abadie, Diamond & Hainmueller (2010). The estimated effect grows over the post-treatment period, from +1.77 percentage points in 2006 to +15.93 by 2022, reaching significance on the standardized p-value ($p = 0.042$) only in the final years.

Robustness check:

Table 6: SDID donor country reform status (male 60–64).

Reform status (during study)	Countries	Combined weight
Major male 60–64 reform	Denmark, France, Germany, Italy, Sweden, Austria, Greece	46.9%
No major male 60–64 reform	Australia, Belgium, Canada, Finland, Ireland, Netherlands, New Zealand, Portugal, Spain, Switzerland, USA	53.1%

Note: Reform status based on whether the country implemented a major reform directly raising the male pension eligibility age for ages 60–64 during the study period. About 47% of donor weight comes from countries that themselves raised the male 60–64 eligibility age during the study period, which motivates the robustness checks that exclude them.

As Table 6 shows, approximately 46.9% of donor weight in the main specification comes from countries that also implemented major reforms raising the male pension eligibility age for ages 60–64 (Austria, Denmark, France, Germany, Greece, Italy, and Sweden), with the remaining 53.1% from non-reforming countries. To ensure the main results are not driven by contamination from these concurrent reforms, the SDID is re-estimated excluding all reforming donors, leaving 12 countries (11 non-reforming donor countries plus Japan). A further specification restricts the donor pool to five countries with no

pension reform during the study period (Belgium, Canada, Ireland, New Zealand, and Switzerland), leaving 6 countries, including Japan.

Table 7: SDID Robustness Test.

Specification	2001 ATT	p	2006 ATT	p	Swing
19 countries	-8.34	0.210	+2.50	0.752	+10.84
12 countries	-6.13	0.472	+4.72	0.655	+10.85
6 countries	-4.90	0.322	+3.87	0.454	+8.77

Note: The swing is calculated by subtracting 2001 ATT from 2006 ATT. The countries include Japan. The negative-to-positive sign flip persists across all three specifications, ranging from 8.8 to 10.9 percentage points, suggesting that the pattern is not driven by concurrent reforms in donor countries.

Table 7 reports these specifications. The results across all three specifications are consistent. Each yields negative ATT when the treatment year is set at 2001 and positive ATT when shifted to 2006, with the sign-flip ranging from approximately 8.8 to 10.9 percentage points. This consistency across progressively restricted donor pools indicates that the main findings are not driven by contamination from concurrent reforms in donor countries.

Interpretation and Policy Implications:

Both synthetic control and synthetic difference-in-differences yield no detectable effect on male LFPR for ages 60–64 with treatment year set as the 2001 reform. When the treatment year is moved to 2006, the 27-country synthetic control shows clear divergence between counterfactual and actual Japan – a 15.93 percentage point difference. And as for the SDID, neither estimate is statistically significant on its own. This pattern is consistent with the hypothesis that supply-side pension incentives require complementary demand-side employment policies to translate into gains in the labor force participation rate. This is expected given Section 3's prediction: Japan's low implicit tax force and LFPR already close to the health capacity ceiling meant there was little room for the reform to produce a detectable effect.

As for the underlying reasons behind the null and negative results for 2001, this paper proposes three main factors. First, the low implicit tax force: Japan's tax force was 160%, meaning Japanese workers were already facing very little penalty for working past the Early Retirement Age. Raising the eligibility threshold further reduced the already-small implicit tax; the change in incentive was modest. Second, the ceiling effect: the estimated health capacity ceiling is 80%, and Japan's pre-reform LFPR of 71 – 76% was already close to it. There was comparatively little room for participation to rise compared to other countries, where LFPR among 60–64 year-olds was 15 – 20%. Third, the demand-side constraint. The majority of corporations in Japan had their mandatory retirement age remain at 60 after the 2001 pension reform. This led to a phenomenon where even if workers were incentivized to keep working, employers could and did force them out. Supply-side incentives did not yield expected benefits as there was insufficient demand to absorb the increased supply.

The 2006 Elderly Employment Stabilization Law removed the demand-side constraint by requiring firms to offer re-employment contracts until the pension eligibility age. Workers

who were incentivized by the pension reform to keep working after 60 now had the legal mechanism to do so. The supply-side push could now operate alongside the demand-side pull, and this is supported by Figure 3, which shows the year-by-year SC results. Japan underperformed the synthetic control from 2001 to 2006, then crossed above it beginning in 2008, which is consistent with the expected effect of the 2006 law.

Kondo and Shigeoka reached the same conclusion as this paper using within-country difference-in-differences,⁷ showing that the 2006 law increased employment and the pension reform's effect was larger when combined with demand-side enforcement. This paper provides additional suggestive evidence for this finding using cross-country SC and SDID, which are entirely different methods, and yet the same conclusion is reached. The convergence of within-country and cross-country evidence strengthens the credibility of the hypothesis that pension reform requires accompanying employment legislation. Additionally, the NBER International Social Security project noted that most developed countries had their pension systems reformed during the study period, reducing implicit taxes and increasing older worker employment.¹¹ This suggests that the donor countries in this study were also changing, which helps explain the null result for 2001; Japan's LFPR may have improved, but so did the counterfactual.

Many developed countries are currently raising pension eligibility ages, such as the UK, Australia, and the Netherlands. This finding has implications for other countries: raising the pension age alone may not work if employers can still force workers out at 60. Governments should pair pension reform with employment legislation that requires firms to retain or rehire older workers.

There are several limitations to this paper. Most crucially, neither the 2001 nor the 2006 ATT is statistically significant at the 5% level on the SDID model, which limits the findings to suggestive patterns rather than confirmed effects. The donor pool is comparatively small, which limits statistical power. Among the countries in the donor pool, some also reformed their pension systems during the study period, which may bias the ATT toward zero by contaminating the "no treatment" assumption. Additionally, only approximately half of the 60–64 cohort in Japan were directly affected by the reform, as the eligibility age increase applied to EPI members (private-sector employees), while NPI-only members, such as the self-employed, already faced an eligibility age of 65. The outcome variable, however, captures aggregate male LFPR for the entire 60–64 age group, which may dilute the estimated treatment effect.

■ Conclusion

This paper used synthetic control and synthetic difference-in-differences to estimate the effect of Japan's 2001 pension reform and 2006 Elderly Employment Stabilization Law on male labor force participation for ages 60–64. The 2001 treatment yielded no statistically significant effect across SC and SDID. However, when the treatment year is moved to 2006, SC shows a 15.93 percentage point increase. This pat-

tern suggests that pension reform alone was insufficient and required the 2006 Elderly Employment Stabilization Law to translate into labor force participation gains. This provides suggestive evidence that countries currently raising pension eligibility age should pair reforms with employer retention mandates or complementary employment legislation.

This paper is the first to use cross-country evaluation on this combination of policy, and it reaches the same conclusion as Kondo & Shigeoka's within-country work. This strengthens the credibility of the finding. For countries that are presently raising or considering raising pension eligibility age to adjust their pension system, the implication is that pension reform without accompanying employer retention mandates risks the same gap issue Japan experienced between 2001 and 2006.

This research is limited by the comparatively small donor pool and by contamination from concurrent pension reforms in donor countries, both of which reduce statistical power. Additionally, only EPI members within the 60–64 cohort were directly affected by the reform, while the outcome variable captures the entire age group, potentially diluting the estimated effect. Moreover, as the reform is phased in, the effect is likely to be diluted over time. Further research using cross-country individual-level data or a larger donor pool could provide more precise estimates of pension reform's effect on elderly employment. A further direction is whether the 2006 law increased elderly participation at the cost of employment quality, as reemployed workers typically accepted substantially lower wages than in their pre-retirement positions. Moreover, the SDID model can yield more accurate results when fed with micro-level data, which presently cannot be obtained from the authorities by me. We should note that because of this, this essay's SDID result should be interpreted with caution.

■ Acknowledgments

I would like to express my gratitude to Professor Jose Maximo for teaching me how economic theory applies to policy research and for guiding my understanding and utilization of econometrics, bringing economics from paper to real-life application.

■ References

1. World Economic Forum. *Japan's aging population: The implications for its economy*. <https://www.weforum.org/stories/2023/09/elderly-oldest-population-world-japan/> (accessed 2026-02-18).
2. Nippon.com. *Births in Japan Fall Below 700,000 for the First Time in 2024*. <https://www.nippon.com/en/japan-data/h02429/> (accessed 2026-02-11).
3. OECD. *Fertility rates (indicator)*. <https://www.oecd.org/en/data/indicators/fertility-rates.html> (accessed 2026-02-11).
4. Japan Times. *Japan's Average Life Expectancy in 2024 Almost Unchanged from 2023, Data Shows*. 2025. <https://www.japantimes.co.jp/news/2025/07/25/japan/science-health/life-expectancy-average-flat/> (accessed 2026-02-18).
5. Ministry of Health, Labor and Welfare. *The Point of the Pension Plan*; Government of Japan, 2010. https://www.mhlw.go.jp/english/org/policy/dl/p36-37_1.pdf (accessed 2026-02-18).
6. Kondo, A. Effects of Increased Elderly Employment on Other Workers' Employment and the Elderly's Earnings in Japan. *IZA*

- J. Labor Policy* **2016**, 5 (2), 1–23. <https://doi.org/10.1186/s40173-016-0063-z>.
7. Kondo, A.; Shigeoka, H. The Effectiveness of Demand-Side Government Intervention to Promote Elderly Employment: Evidence from Japan. *ILR Rev.* **2017**, 70 (4), 1008–1036. <https://doi.org/10.1177/0019793916676490>.
 8. Shimizutani, S.; Oshio, T. Revisiting the Labor Supply Effect of Social Security Earnings Test: New Evidence from Its Elimination and Reinstatement in Japan. *Jpn. World Econ.* **2013**, 28, 99–111. <https://doi.org/10.1016/j.japwor.2013.08.003>.
 9. Okumura, T.; Usui, E. Effect of Pension Reform on Pension-Benefit Expectations and Savings Decisions in Japan. *RIETI Discuss. Pap. Ser.* **2011**, No. 11-E-065.
 10. U.S. Social Security Administration. *Social Security Programs Throughout the World: Asia and the Pacific, 2018 – Japan*. <https://www.ssa.gov/policy/docs/progdesc/ssptw/2018-2019/asia/japan.html> (accessed 2026-02-18).
 11. Coile, C.; Wise, D.; Börsch-Supan, A.; Gruber, J.; Milligan, K.; Woodbury, R. Social Security and Retirement around the World: Lessons from a Long-Term Collaboration. *J. Pension Econ. Finance* **2025**, 24 (1), 8–30. <https://doi.org/10.1017/S1474747223000215>.
 12. Gruber, J.; Wise, D. A. *Social Security Programs and Retirement around the World: Micro-Estimation*; University of Chicago Press: Chicago, IL, 2004.
 13. World Bank. World Development Indicators, 2024. <https://databank.worldbank.org/source/world-development-indicators> (accessed 2026-02-18).
 14. DiNardo, J.; Fortin, N. M.; Lemieux, T. Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach. *Econometrica* **1996**, 64 (5), 1001–1044. <https://doi.org/10.2307/2171954>.
 15. International Labor Organization. ILOSTAT - International Labour Organization Statistics, 2024. <https://ilostat.ilo.org/> (accessed 2026-02-18).
 16. Kritzer, B. E. Chile's Next Generation Pension Reform. *Soc. Secur. Bull.* **2008**, 68 (2), 69–84.
 17. OECD. *Pensions at a Glance 2007: Public Policies Across OECD Countries*; OECD Publishing: Paris, France, 2007.
 18. Korea Development Institute. *National Pension Reform in South Korea*; Korea Development Institute, 2024. https://www.kdi.re.kr/eng/research/focusView?pub_no=18361 (accessed 2026-02-18).
 19. Abadie, A.; Diamond, A.; Hainmueller, J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *J. Am. Stat. Assoc.* **2010**, 105 (490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>.
 20. Arkhangelsky, D.; Athey, S.; Hirshberg, D. A.; Imbens, G. W.; Wager, S. Synthetic Difference-in-Differences. *Am. Econ. Rev.* **2021**, 111 (12), 4088–4118. <https://doi.org/10.1257/aer.20190159>.

■ Author

Terry S. H. Lo is presently a Year 12 student at Cardiff Sixth Form College in the United Kingdom. He is an inquisitive student interested in understanding the mechanisms behind economic phenomena and applying economic theory to real-world policy questions. He will continue to navigate the vast landscape of economic academia.