I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14469

Pension Incentives and Labor Supply: Evidence from the Introduction of Universal Old-Age Assistance in the UK

Matthias Giesecke
Philipp Jaeger

JUNE 2021

DISCUSSION PAPER SERIES

IZA DP No. 14469

Pension Incentives and Labor Supply: Evidence from the Introduction of Universal Old-Age Assistance in the UK

Matthias Giesecke
Federal Ministry of Finance, RWI and IZA

Philipp Jaeger
RWI

JUNE 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Pension Incentives and Labor Supply: Evidence from the Introduction of Universal Old-Age Assistance in the UK*

We study the labor supply implications of the Old-Age Pension Act (OPA) of 1908, which, for the first time, provided pensions to older people in the UK. Using recently released census data covering the entire population, we exploit variation at the newly created age-based eligibility threshold. Our results show a considerable and abrupt decline in labor force participation of 6.0 percentage points (13%) when older workers reach the eligibility age of 70. To mitigate the impact of population aging today, pension reforms aimed at increasing elderly labor supply, however, have to induce much larger behavioral responses than the OPA.

JEL Classification: D61, H21, H55, J14, J22, J26

Keywords: old-age assistance, labor supply, retirement, regression discontinuity design, equity-efficiency trade-off

Corresponding author:
Matthias Giesecke
RWI
Hohenzollernstr. 1-3
45128 Essen
Germany
E-mail: matthias.giesecke@rwi-essen.de

* We are grateful to Ronald Bachmann, Stefan Bauernschuster, Sebastian Braun, David Card, Fabian Dehos, Erik Hornung, Robin Jessen, Anica Kramer, Christian Proano, Laura Salisbury, Han Ye and the participants of several conferences and seminars for valuable comments and insightful discussions. Johanna Muffert provided excellent research assistance.

1 Introduction

Many pension programs intend to prevent poverty through means-tested income transfers to the elderly. A large body of literature has documented that labor supply responds to marginal changes in existing pension schemes. Little is known, however, on the labor supply effects of introducing universal old-age assistance programs for the first time.

This article assesses the labor supply implications of the Old-Age Pension Act (OPA) of 1908 that established universal means-tested pensions in the UK at the expense of higher taxes for top income earners. In contrast to most of the existing literature, the unique historical setting allows us to estimate the labor supply response when moving from essentially no program to an old-age assistance scheme that covered a large share of the elderly population.[1] Using recently released full-count population data from the UK census in 1911, we isolate the causal effect of the program along the lines of an age-based eligibility threshold that has been introduced by the OPA at the age of 70.[2] This threshold induces a discontinuity in the retirement probability that we use to identify changes in the labor force participation (LFP) rate by adopting a regression discontinuity (RD) design.

The historical setting, besides being interesting in its own right, also involves several other features that strengthen the empirical analysis. Enacted in 1908 and effectively implemented in 1909, the OPA was the largest means-tested old-age assistance program at that time[3] and served as an archetype for public old-age assistance in the US.[4] Although the UK was a large industrialized economy when the OPA was introduced, it still lacked a comprehensive welfare

[1] In 1911, 57.5% of people who reached the eligibility age of 70 did receive a pension that was administered through the OPA.

[2] Despite the low expectancy at birth, the UK had a sizable elderly population at that time (more than one million people were 70 and older in England and Wales alone). Adjusting for child mortality shows that about 35% of the birth cohort 1841 reached age 70 in 1911 (life tables Human Mortality Database, 2018). Conditional on having celebrated the 70th birthday, this cohort had a remaining lifetime of 9 years on average.

[3] Germany already introduced a public pension scheme in 1891, but pensions were based on contributions and not means-tested. Belgium, France, Denmark, New Zealand and parts of Australia had means-tested old-age assistance programs in place before 1909. However, these programs were smaller with respect to the absolute number of recipients or the benefit level.

[4] After the introduction of the OPA, public pensions, which yet did not exist except for veterans, also made the agenda in several US states Chamberlain and Pierson (1924). Most of the following pension legislation was based on a provision known as "the standard bill" Bureau of Labor Statistics (1929) which was remarkably similar to the OPA and encompassed means-tested, non-contributory pensions starting with the age of 70.

system. Private pension schemes and other government programs that addressed older workers were either small and uncommon or did not exist. Furthermore, program substitution towards unemployment or health insurance was impossible because these programs became effective in 1912/1913 and thus after the implementation of the OPA (1909) and data collection (1911).

Previous studies have suggested that the introduction of the OPA had little impact on labor supply (Johnson, 1994; Costa, 1998). This conjecture is based on aggregate labor force participation (LFP) rates from British censuses, suggesting that the downward trend in LFP rates among older men in the UK did not accelerate after the introduction of the OPA. Using more disaggregated data, we provide evidence that LFP rates of older men and women did decline promptly and considerably as a direct consequence of the introduction of the OPA.

Our main findings indicate that, in absolute terms, the LFP decline amounts to 6.0 percentage points at the age of 70 just after the implementation of the OPA. Relative to a participation rate of 46% at age 69, the LFP rate thus declines by 13% when moving marginally above the age-based eligibility threshold. The LFP drop is predominantly driven by reductions in work activity and much less by lower unemployment. Our estimates are consistent with substitution effects and constrained income effects. The means test puts an implicit tax on labor earnings (i.e. substitution effect), which makes labor (relative to leisure) less attractive and results in retirement bunching at age 70. At the same time, liquidity constraints, myopic behavior, or uncertainty about future pension policies might prevent that the transfer component of the pension (i.e. income effect) leads to retirement prior to the eligibility age. These market imperfections are plausible explanations for income effects only to materialize in a discrete downward jump of labor supply at the eligibility age (and not before).

Labor supply responses are considerably larger for people in low-earnings occupations compared to high-earnings occupations. We document that the LFP reduction of 15.3% in the lowest occupational earnings quartile more than doubles the LFP decline of 7.2% in the highest quartile. This finding is consistent with the theoretical prediction that people with low earnings should respond more strongly to means-tested cash transfers because they have to forgo less earnings to pass the means test than people with high earnings. Our results also indicate that the LFP decline is smaller for older workers with close family ties in comparison to individuals without. This finding is in line with the hypothesis that private transfers from children, spouses, or other household members served as old-age insurance when public pensions were absent. We further provide evidence for joint retirement decisions of couples. We document a significant LFP drop of 3.3 percentage points among eligible men when their spouse also reaches the eligibility age.

The OPA was most likely welfare enhancing, despite the clear labor supply decline. The OPA generated only modest efficiency losses, because the behavioral responses to the pensions were moderate relative to the scale of the program. Moreover fiscal externalities were small because of the low tax rates and limited other social spending at the time. In contrast, the degree of redistribution, and hence the equity gain, was immense. The policy redistributed from the absolute top to the bottom of the earnings distribution in a time of high income inequality.[5]

This paper adds to earlier studies on labor supply responses to pension reforms. The literature has mostly focused on marginal changes in existing pension systems (e.g. Krueger and Pischke, 1992; Börsch-Supan, 2000; Mastrobuoni, 2009; Liebman et al., 2009; Brown, 2013; Danzer, 2013; Atalay and Barrett, 2015; Manoli and Weber, 2016). We extend this strand of the literature by studying a first-time introduction of an old-age assistance system, thus investigating the change from no program to universal coverage. Our estimates quantify the immediate labor supply effects of the OPA, starting from a benefit level of zero without having to rely on extrapolations. Such estimates are informative for the introduction of old-age assistance systems in today's developing countries and also serve as a benchmark for marginal changes in contemporaneous pension programs. A simple back-of-the-envelope calculation reveals that future pension reforms that aim to stabilize the retiree-to-worker ratio have to induce much larger behavioral responses than the OPA.

The only existing evidence on the labor supply effects of introducing public pensions for the first time stems from studies on the US Social Security Act of 1935 (Parsons, 1991; Friedberg, 1999; Fetter and Lockwood, 2018). In contrast to the US program that was introduced during the turbulent times of the Great Depression, old-age assistance in the UK was established in a period of stable economic conditions which ensures that our results are not confounded by

[5] In 1911, 8.4% of total income (before taxes) in the UK went to the top 0.05% compared to 3.4% in 2000 (Atkinson, 2007).

macroeconomic shocks or the massive expansion of other social spending. Moreover, the institutional setting in the UK allows for highly credible and transparent identification based solely on an age RD. Our research design builds on many recent studies that have used age-based eligibility thresholds to identify policy-relevant effects (see e.g. Card et al., 2008; Battistin et al., 2009; Carpenter and Dobkin, 2009; Card et al., 2009; Anderson et al., 2012, 2014; Carpenter and Dobkin, 2015, 2017; Fitzpatrick and Moore, 2018).

The remainder of this paper is structured as follows. Section 2 provides historical and institutional details on the old-age assistance program in the UK and how its introduction creates exogenous variation that we use for identification. Section 3 outlines the research design and describes the data. Section 4 presents results, sensitivity checks and falsification tests. Section 5 discusses mechanisms, welfare implications and policy implications of the program. Section 6 concludes.

2 Historical Background and Institutional Details

The Old-Age Pension Act of 1908 (OPA) introduced means-tested, non-contributory minimum pensions for British citizens financed by the central government. The OPA was a major social policy intervention and the first one to specifically target the elderly in a time of very limited social protection. The law was debated in the British Parliament in May 1908, passed through in August 1908 and the first pensions were eventually paid out in January 1909. At that time, neither unemployment nor health insurance existed because both of these programs only became effective in 1912/1913.

Given that pensions were means-tested, the coverage of the OPA was astonishingly high. In 1911, almost 60% of people who had reached the eligibility age were granted a pension in England and Wales (613,873 out of 1,068,486 according to the Department of Labour Statistics, 1915, p. 184). The vast majority of pension recipients (about 93% in 1911) received the maximum pension of 5 *shillings* per week. According to Feinstein (1990), this amounted to approximately 22% of average earnings at the time.

The OPA was a response to the perceived inadequacy of the existing poor relief system that provided only very basic protection and involved considerable sanctions (including the loss of voting rights) and social stigma as well as the requirement to work in a workhouse unless the person could prove to be sufficiently unfit. The newly introduced pensions were not only less restrictive but also involved more generous benefits and thus considerably more older people applied for them.[6] In contrast to the poor law, which was administered and financed at the local level giving local authorities a lot of discretion in the assignment of financial aid, the OPA was enacted as a nation-wide right for older workers who met the specified pension eligibility criteria. The introduction of the pension reduced the number of poor relief recipients among the elderly substantially, however, it had no apparent effect on poor relief support for people below age 70.[7]

To reduce social stigma, people could apply for pensions at the local post office. Decisions were made by a pension officer appointed by the treasury and a local pension committee. Pension eligibility was mainly based on two criteria: age and inadequate means.[8]

Older workers only became eligible when reaching the age of 70. The original proposal for the reform, dating back to 1899, recommended a retirement age of 65, which would have been more in line with the retirement rules in the few pre-existing pension schemes that typically

[6] In 1906, 23.3% of people aged 70+ in England received some form of poor relief (Boyer, 2016, table 13) compared to 57.5% of people 70+ that received a pension in England and Wales in 1911. In contrast to the pension, poor relief was often tied to a residence in a workhouse. For instance, in 1908, more than 60% of elderly poor relief recipients in London were relieved in workhouses (Boyer, 2016, table 7). For people that were relieved outside of workhouses, the poor relief payments were typically much smaller than the newly introduced pension. According to Pugh (2002), weekly poor relief was on average 2 *shillings* and 6 *pence* and hence only 50% of the pension.

[7] Neither spending per poor relief recipient nor the number of poor relief recipients below the age of 70 appear to be substantially affected by the OPA. According to Department of Labour Statistics (1913, p.328-329), mean annual expenditure per person in receipt of poor relief (outside of a workhouse) has increased by less than 3% from 1908 to 1912 (from £7 and 1 *shilling* to £7 and 5 *shillings*). At the same time, the total number of people receiving poor relief in England and Wales declined by 122,320 (from 772,346 to 650,626). The decline almost perfectly matches the number of people formerly relying on poor relief that became old age-pensioners in the same time period (122,415). Moreover, the share of able-bodied poor relief recipients relative to the total population, which can serve as a proxy for young poor relief recipients, remained constant.

[8] Moreover, the law included a residency requirement of 20 years. If the claimant had satisfied all of the the eligibility criteria, he could still be disqualified due to the following reasons. First, receiving poor relief or having received poor relief any time between January 1908 and December 1910. Second, habitually failing to work according to her ability. And third, being detained in a lunatic asylum, or in any place as a recipient of poor relief or a criminal lunatic or being in jail (or ordered to be imprisoned) less than ten years ago.

specified an age between 60 and 65.[9] However, the original suggestion was considered too expensive. Proving age was straightforward for pension claimants in England and Wales because birth registers (and thus birth certificates) existed at least since 1837. Verification was more difficult in Scotland and especially challenging in Ireland, making rejections based on the age criterion far more frequent in Ireland (Old Age Pensions Committee, 1919). Given the low life expectancy at birth (below 50 for males in 1911), a retirement age of 70 seems high by today's standards. The low life expectancy, however, was mainly driven by high infant mortality. Once reaching the age of 70 in 1911 (birth cohort 1841), people could expect to live another 9 years on average (men: 8.5 yrs., women: 9.5 yrs. Human Mortality Database, 2018).

Eligibility was also conditional on a means test. Claimants had to prove that their annual means were below 630 *shillings* (54% of average annual earnings at that time) to receive any pension. To become eligible for the maximum pension, an income of less than 420 *shillings* (36% of average annual earnings) was required.[10] Means were calculated based on labor income, family transfers as well as returns on property, including hypothetical returns and the rental value of living in one's own house. Spousal income was also considered explicitly.[11] The law also prescribed that if individuals intentionally deprived themselves of resources, the value of these resources would still be included in the calculation of the annual means. However, one should not overestimate the strictness of the means test, since means were typically self-reported.[12] Moreover, the local pension committees – the most important decision making authority – were inclined to generosity (Pugh, 2002) and thus most applications (around 90%)

[9] Small-sized pension schemes within the boundaries of the respective firm existed previous to the large-scale introduction of the OPA. However, these pensions were rather informal and discretionary. More formalized schemes only existed in very few larger firms, for example in the railway industry, or the public sector. By far, these pensions did not reach the universal coverage rate that was introduced through the OPA (see Thane, 2000, for details on pre-existing pensions in the UK).

[10] In theory, pensions were granted based on a gliding scale as depicted in appendix figure B.1. However, the overwhelming majority (93% in 1911) received the maximum pension of 260 *shillings* per year.

[11] Spouses living together in one household had to fulfill two criteria: 1) Their own income and 2) the per person income of the couple had to fall below the threshold. An example provided in Casson (1908) clarifies this rule: imagine a married couple. The man earns 800 *shillings* and the woman 200 *shillings*. In this case, the woman would be eligible for a pension and has to claim an income of 500 *shillings* ((800+200)/2). The man, however, would not be eligible as his own income of 800 *shillings* exceeds the threshold.

[12] Newspaper articles suggest that fraudulent claims based on misreported means, however, were indeed (at least sometimes) legally prosecuted, indicating that the means test was not simply a paper tiger.

were approved (The Times, 1909; Old Age Pensions Committee, 1919)[13] and only 3.5% of pension decisions were appealed in 1913 (Pugh, 2002).

The government expenses for the pension were therefore considerable, despite the relatively high retirement age. In the budget year 1911, £6.3 million were spent on old-age pensions in England and Wales. For the entire UK, old-age pension expenditures amounted to £9.8 million, making pensions one of the largest single spending item (5.7% of overall expenditures in the budget of 1911, House of Commons, 1911). Pensions were financed almost exclusively by taxes on high income earners (roughly the top 1% in the UK).[14] For this purpose, new progressive tax brackets were introduced (House of Commons, 1911): The marginal tax rate for people with annual earnings of £2,000 - £3,000 increased by around 1.2 percentage points (to 5.0%), for £3,000 - £5,000 by around 2 percentage points (to 5.8%) and above £5,000[15] by around 4.5 percentage points (to 8.3%). The marginal tax rate for the majority of tax payers earning between £160 and £2000 was not increased and remained at 3.8%.[16] Most people did not pay any income tax because their earnings were below the taxation threshold of £160 (2.7 times the mean earnings) and were also not affected by income tax changes.[17]

3 Empirical Strategy and Data

3.1 Exogenous Variation and Research Design

To identify the causal effect of pension availability on LFP, we take advantage of the age cutoff that was introduced by the OPA. Pension eligibility at age 70 creates a discontinuity in the local environment between age 69 and 70. Along the lines of this age threshold, we adopt an RD design with the age as assignment variable. The identifying assumption is that the

[13] The main reasons for disapproval were the inability to prove the individual age (most relevant in Ireland), a failed means test or receiving poor relief.

[14] Other less important components of the financing system were death duties as a type of inheritance tax and stamp duties on investments. For details on the financing sources of the OPA, see Murray (2009).

[15] Earnings above £5,000 constituted the top 0.05% of earnings, according to Atkinson (2005).

[16] This group included about 75% of tax payers, e.g. 750,000 out of 1,000,000 (Murray, 2009).

[17] At that time, only one million people paid income taxes in the UK (Murray, 2009) compared to more than 36 million people living in England and Wales alone.

outcome of interest, LFP, would evolve smoothly between age 69 and 70 if the OPA had not been introduced. Any discontinuous jump of the outcome at the eligibility cutoff can be attributed to the availability of the pension if other programs did not affect LFP at the respective age.

3.2 Estimation

The observable outcome y_a is an indicator of LFP that takes the value one if the individual is in the labor force at age a and zero otherwise. We thus estimate the equation

$$y_a = \beta_0 + \beta_1 \mathbb{1}(a \geq 70) + \beta_2 f(a) + \varepsilon_a \tag{1}$$

where the coefficient of primary interest, β_1, measures the percentage point difference in LFP, comparing the share of people in the labor force marginally above the age cutoff (age 70) to the respective share marginally below the age cutoff (age 69). To account for the possibility of a functional relationship between the outcome LFP and the assignment variable age, the function $f(a)$, which is allowed to vary on either side of the age cutoff, not only includes age linearly but also as a second order polynomial. However, graphical evidence suggests that the age-LFP relationship is essentially linear close to the age-cutoff.

To show that our results are exceptionally robust against changes in the specification, we implement several alternative estimation procedures that are common in the RD literature. We extend the baseline estimation framework with uniform weighting to more flexible local non-parametric estimates that put more weight on observations close to the cutoff (triangular kernel weighting). We also present bias-corrected point estimates as suggested by Calonico et al. (2014a,b) and provide detailed results on how the estimates differ by bandwidth choice and the order of the polynomial. Finally, we show that using alternative types of standard errors does not change our conclusions (see section 4.3).

3.3 Data and Summary Statistics

The analysis relies on full-count individual level census data for three decennial UK census waves collected in the spring of 1891, 1901 and 1911. The dataset is a recent release by the Integrated Census Microdata (I-CeM) project (Higgs et al., 2013), distributed by Integrated Public Use Microdata Series International (IPUMS International Minnesota Population Center, 2018).[18] We use information for England and Wales, thus excluding Scotland, Ireland and the Channel Islands because data is not available for the other regions at all points in time. Moreover, birth certificates, which substantially reduce age-misreporting, only existed in England and Wales for a sufficiently long period. Finally, we exclude persons with unknown gender (less than 0.1 % of the population) or age (0.2 % of the population).

3.3.1 Dependent Variable

Our definition of the labor force status is based on the gainful employment concept which was used before the UK adopted the current labor force definition in 1961. In contrast to the current definition, which categorizes people based on their activity status (working or seeking work) in a specific reference week, the gainful employment concept derives the labor market status from the occupation of the respondents. In particular, we include people in the labor force (LFP = 1) if they specify an occupation or report to be unemployed.[19] Individuals are considered out of the labor force (LFP = 0) if they report no occupation or that they have retired from a specific occupation.[20] The current definition of LFP and the gainful employment concept are closely related. Costa (1998) constructs participation rates based on the gainful employment concept for the US until the 1990s, showing that the patterns of both series match. Similarly, Johnson (1994) argues that the change of the definition in 1961 did "appear to have

[18] The I-CeM project collaborated with the website findmypast.org to transcribe and harmonize several historical British censuses, encompassing data collected in the years 1851, 1861, 1881, 1891, 1901, and 1911. Recent economic studies that have used selected waves are, for example, Arthi et al. (2021) and Beach and Hanlon (2019).

[19] We also include people that report to be formerly employed in the labor force. Recoding this subgroup as not in the labor force does not affect our results.

[20] Following the census in 1891, retirement was explicitly recognized as a separate category and retirees were not considered economically active anymore (Johnson, 1994), which is arguably consistent with being out of the labor force. We adjust the labor force variable constructed by IPUMS International by defining individuals to be out of the labor force if they state an occupation but add that they have retired already.

had little effect on the enumeration of older workers" (Johnson, 1994, p. 109). Based on this evidence, the two concepts arguably yield very similar patterns of LFP over time. For our empirical analysis, potential differences between the two concepts are of little relevance because the gainful employment concept did not change throughout the time under study (1891 - 1911). Differences only need to be kept in mind when comparing the results to the current LFP concept.

3.3.2 Summary Statistics

Using full-count census data enables us to zoom in directly at the age cutoff. Figure 1 shows the distribution of observations over age for the 1911 census, including 150,293 individuals at age 69 and 140,288 individuals at age 70. Except for minor round number bunching[21] the sample size declines almost steadily with age. While the 1911 census counts more than 400,000 individuals at age 50, the number of individuals drops below 5,000 at age 90.

Summary statistics in table 1 (upper panel) report a considerable decline in LFP between age 69 and 70. The drop totals to 7 percentage points (from 46% to 39%), while differences in other observable characteristics are fairly small. At age 70, individuals are less often married and the share of foreign born individuals is slightly higher. The lower panel in table 1 reports summary statistics for the baseline estimation sample with five age-years below the cutoff (65 - 69, N: 803,208) and above the cutoff (70 - 74, N: 551,100). Including additional age-years naturally leads to a larger differential in mean LFP rates of 48% below the eligibility age and 34% above.

Throughout this study, we examine the labor supply responses jointly for men and women. Since the participation rates of women were generally low, we also present the main results separately for men and women.[22] In general and for both men and women, the decline of

[21]Note, for example, the small dip in the number of observations at age 71. Such spikes in the sample size at round ages ("age-heaping") are common in (historical) survey data. People might round their age either because they are uncertain about their exact age, or because they are innumerate or inattentive. We discuss this issue in more detail in section 4.3.3.

[22]Table 1 indicates that the average LFP rate marginally below the eligibility threshold (age 69) was 46%. However, while only 20% of women were still economically active at that age, the LFP rates of older men were relatively high and totaled to 76% (figure 7). See section 4.2.3 for separate results regarding men and women.

LFP rates over age evolved less steeply than today. We will argue in this paper that one major explanation for this phenomenon was the low coverage by social security and old-age pensions in the early 20th century.

The majority of older workers (almost one third) was active in sectors such as crafts or related trades (table 2). The second largest sector included service workers, followed by skilled agricultural and fishery work and elementary occupations. In contrast, only few older workers earned a living as senior officials or managers, technicians or other professionals. We use occupational information later-on to construct a measure of occupational labor earnings. This allows for estimates of the labor supply response to pension availability in different regions of the occupational earnings distribution.

4 The Labor Supply Effect of the OPA

4.1 Baseline Results

Table 3 reports the decline in labor supply due to the OPA for our preferred RD specification, using a bandwidth of 5 age-years to the left (65 - 69) and to the right (70 - 74) of the age cutoff, a linear age polynomial[23], and uniform weighting on all observations. The estimate shows a precisely measured drop in LFP of 6.0 percentage points when people reach the age-based eligibility threshold. Contrasting the LFP distribution over age separately for the 1901 and the 1911 censuses, figure 2 reveals the striking difference in age-specific LFP patterns before and after the OPA became effective in 1909. Without universal coverage by old-age assistance in 1901, participation rates decline gradually over age. In contrast, we observe an abrupt drop of LFP in 1911 exactly at the age-based eligibility cutoff. The sudden drop only appears at the eligibility age and only after the OPA has been introduced. Since no other social security programs existed that would induce LFP changes between age 69 and 70, the LFP reduction is caused by the availability of the OPA.

[23] Graphical evidence in figure B.2 suggests that linear and quadratic age polynomials reproduce the LFP-age relationship very similarly and thus, for brevity, we report the linear specification for all subsequent RD specifications including placebo tests and robustness checks. Throughout, the results are very similar for the quadratic polynomial and are available from the authors upon request.

The estimated decline is sizable both in absolute and relative terms. Departing from a participation rate of 46% at age 69, the estimated absolute decline of 6.0 percentage points translates into a relative decline of 13%. Given the scale of the program and the limited social security net at that time, the substantial decline in participation rates is not surprising.

The considerable negative impact of the OPA on LFP rates of older workers, however, can be easily overlooked in aggregated labor market trends. As noted by Johnson (1994) and Costa (1998), the decline of LFP rates for the aggregate of older men in the UK has not accelerated after the OPA was implemented in 1909. Using a simple difference-in-differences (DiD) representation (table 4, figure 3)[24], we argue that simply extrapolating past aggregate LFP trends, however, does not produce a valid counterfactual scenario. Figure 3 shows that participation rates of people slightly below the eligibility age decreased between 1891 and 1901 but increased between 1901 and 1911. Thus, we believe that the pace of the LFP decline for older workers would have slowed down between 1901 and 1911 without the OPA. We extend the DiD framework in section 4.3.4 to argue that there is little evidence for anticipation effects, i.e. people adjusting their LFP before age 70, lending credibility to the argument that people slightly below the age of 70 are indeed a valid control group for the people slightly above age 70.

Our results are consistent with those of Fetter and Lockwood (2018), who study the introduction of the old-age assistance program in the US. They find that the labor supply of men aged 65 to 74 declined by 8.5 percentage points between 1930 and 1940 as a consequence of the introduction of old-age assistance in 1935. Our estimated LFP decline of 6.0 percentage points is slightly smaller in magnitude, which is mostly due to the inclusion of women in our sample (for more information on gender differences see section 4.2.3). In contrast to the US

[24] For this exercise, we add the census wave of 1891 in addition to 1901 and 1911, which allows us to graphically test the common trend assumption. For the two age groups at the eligibility threshold set by the OPA, the graph indicates a common trend in the pre-treatment period between 1891 and 1901, showing that LFP rates of the two groups move in tandem between 1891 and 1901. The common trend is stable for neighboring ages below the eligibility threshold (65-69) and above (70-74), as is evident from figure B.3 in the appendix. After that, between 1901 and 1911, the two age groups diverge in terms of LFP. The DiD estimate quantifies the LFP differential to range between 4.0 to 4.5 percentage points difference between older workers aged 69 and 70 (table 4). It eliminates any time-constant unoberservables that would confound estimating the impact of the introduction of the OPA on LFP. The DiD is robust against including any available observable characteristics such as number of own children in the household, household size, and indicators for being married, foreign born, disabled, and region (covering the 53 counties in England and Wales).

case, where almost half of the decline was explained by exits from work relief programs and unemployment, direct transitions from work to retirement were more prevalent in the UK.

To examine the labor supply response of the population that retires directly from work, we now disregard the unemployed[25] and deviate from the main LFP measure by setting LFP = 1 only for those who are actually working (and LFP = 0 for everyone else). We assume that those individuals in the data who report an occupation are working and that those who report to be unemployed are not working.[26] This exercise shows that almost the entire decline in LFP is driven by people who stop working (figure 4). The LFP rate among active workers declines significantly by 5.7 percentage points at the age-based eligibility threshold (table 5, column 1). When setting LFP = 1 for those in the labor force who are unemployed, the drop is small (table 5, column 2). The declining LFP rates are hence strongly driven by older workers who directly retire from work.

4.2 Heterogeneity Analysis

4.2.1 Labor Supply Effects Across the Occupational Earnings Distribution

A key theoretical implication of the earnings test is that labor supply responses will differ across the labor earnings distribution. Substituting labor earnings with old-age assistance becomes increasingly costly as labor earnings increase, because people with high-earnings have to forgo more income (and consumption opportunities) to become eligible.

[25]We consider people to be unemployed if they state to be formerly employed (1.2% of the population aged 50-90) or unemployed (0.1% of the population aged 50-90).

[26]People stating an occupation are coded as working, according to the I-CeM data codebook (Higgs et al., 2013). However, as noted in the codebook the "boundaries between the groups [people who are working or being inactive for different reasons] are rather fluid", p.203. In addition, the 1911 census likely undercounts unemployment, because our own calculation classifies only 0.5% of working age men (15 to 69) as unemployed or formerly employed while a sample of trade union members quantifies the unemployment rate in England and Wales at the time of the census (April 1911) to be rather 2.8% (Department of Labour Statistics, 1913, p.6). Although this may indicate a slight overestimation of the share of people working, the total unemployment rate was extremely low at the time such that our result –most people who retired were previously at least occasionally working– is very unlikely to be driven by the mis-classification of unemployed workers as working.

We test this prediction using occupational earnings as a proxy for individual earnings, which are not reported in the 1911 census.[27] Figure 5 shows that negative labor supply responses are more pronounced at the lower end of the occupational earnings distribution than at the top. In the lowest earnings quartile, LFP declines by 15.3% and thus more than twice compared to the highest earning quartile (7.2%). The labor supply response gap between high- and low-earners based on individual earnings data might be even bigger, since the significant decline in the highest occupational earnings group could arguably be driven by low-earnings people working in high-earnings occupations.

4.2.2 Family Background and Old-Age Insurance

The family can function as old-age insurance, especially when individuals are credit constrained and do not have access to social security. The insurance argument has mainly been pointed out with regard to children (Leibenstein, 1957; Caldwell, 1982; Cain, 1983; Boldrin and Jones, 2002), but public old-age assistance can replace any type of family-related transfers (e.g. from spouses or other related household members). This mechanism was recognized during the legislation process of the OPA and, after some debates, it was finally decided that voluntary family transfers must be included in the calculation of the annual means of a pension claimant (Casson, 1908).

Given that private transfers are considered in the means test, family ties might affect eligibility and the corresponding labor supply response to the OPA. Older workers who receive private transfers (e.g. from family members) might either not pass the means test or have already used the family transfers to retire before the age of 70. Consequently, their labor supply response at the eligibility age is expected to be less pronounced. In contrast, individuals who do not receive private transfers have to work until pension benefits become available (if they lack savings) and thus are expected to react more strongly when they reach the eligibility age.

Although we do not have information on within-family cash transfers, we can test three dimensions of family ties based on existing living arrangements and marital status. First, we

[27] We construct occupational earnings based on the reported occupational title on the five-digit occupational HISCO level in the UK 1911 census as well earnings data by occupation from the US census in 1950. See appendix C for details including the resulting occupational earnings distribution in figure B.4.

proxy inter-generational transfers from children by distinguishing older workers with and without own children in the household.[28] Second, we use marital status to proxy transfers in spousal relationships. And third, we compare single versus multiple person households as a more general measure for transfers from closely related individuals.

Table 6 reports RD estimates from samples that are stratified by individuals with and without own children in the household. These estimates indicate that older workers without own children in the household show stronger declines in LFP rates: at the age-based eligibility threshold, LFP falls by 6.9 percentage points. In contrast, LFP declines by only 5.1 percentage points among older workers with children (z-statistic of significant different differences= 2.0). Graphical evidence in figure 6 supports the estimation results. The finding of stronger LFP reactions among older workers without children is consistent with the hypothesis that children served as a type of old-age insurance before social security systems existed .

The insurance argument is particularly salient when looking at solitaire individuals. Consistent with this view, the labor supply reductions among non-married individuals are larger compared to the sub-sample of married individuals (table 6, figure 6) although these estimates do not significantly differ (z-statistic= 0.6). Individuals living completely on their own (single households) respond very strongly to the pension incentive. Individuals living alone show large LFP reductions of 9.5 percentage points, while those living with one or more other persons react significantly less (5.7 percentage points, z-statistic=-3.9).

One concern is that living arrangements correlate with earnings. In fact, we do find that earnings are lower in each of the three groups that show more pronounced labor supply reductions (no children, non-married, and living alone). However, we argue that the results are not entirely driven by earnings gaps for two reasons. First, the occupational earnings gaps are rather small (£6 to £8 on average). Second, the occupational earnings gaps are not correlated with the labor supply response gaps. For instance, we see the biggest occupational earnings gap between married and non-married people, but the smallest LFP response gap. Another concern is that we might condition on outcome variables, since the household composition and marriage incentives could also be affected by the pension. We consider this to be a minor concern since

[28] Note that given the age of their parents almost all children are in working-age already. More than 95% of children are 15 years or older, more than 90% are 20 years or older.

the smoothness analysis in section 4.3.1 suggests that there is little evidence for a substantial change in marriage rates and living arrangements at the cut-off.

4.2.3 Labor Supply Effects by Gender

A particularly robust finding documented in the literature on labor supply is that women have much larger labor supply elasticities than men, especially on the extensive margin of labor supply (see Keane, 2011, for a review). One important difference between men and women, in particular when using data from the early 20th century, is that a large share of women was not part of the labor force based on the prevalent definition. This also holds in the cohorts under study here and is apparent from comparisons of LFP rates between older men and women (figure 7). The figure indicates that LFP rates at the age of 69 were much larger among men (76%) compared to women (20%). RD estimates on separate male and female samples (table 7) show that, in absolute terms, the labor supply responses are larger among men (7.9 percentage points) as compared to women (3.2 percentage points). When placing the smaller female labor supply reduction into the context of their smaller LFP rate, their relative reduction of 16% is much larger compared to men (10%). This result is consistent with larger female labor supply elasticities as documented in the literature.

RD estimates across the occupational earnings distribution also differ between men and women (figure 8). Especially in the second earnings quartile, women show dramatic labor supply reductions of up to 27%. Due to their relatively low occupational earnings[29], women also had low adjustment costs in terms of forgone earnings to pass the means test. This is largely in line with the observation that the female labor supply response is particularly large at the lower end of the female occupational earnings distribution.

4.2.4 Within-Family Spillovers

Many empirical studies suggest that within-family spillovers play an important role in retirement decisions (Blau and Riphahn, 1999; Gustman and Steinmeier, 2000; Baker, 2002; Atalay

[29] The mean occupational earnings of women in 1911 are 41£, which amounts to two thirds of mean earnings among men (60£).

et al., 2019; Garcia-Miralles and Leganza, 2021). The interdependence of spousal retirement behavior might be driven by a correlation of tastes for leisure between partners or the eligibility rules of the pension system that takes spousal income into account.

The UK census allows us to link the labor force status of spouses. We use this information to study whether people respond to the pension eligibility of their spouse. We focus on the impact of female eligibility on male retirement patterns, because LFP rates of married elderly women at the time were extremely low (around 8% in the relevant age range) which leaves little room for additional female retirement. To provide graphical evidence, we first plot male LFP rates against spousal age (instead of own age). Throughout the analysis, we restrict the sample on men that are older than 70, and thus have already reached the eligibility age, to avoid that our results are driven by same age couples reaching the eligibility age at the same time.[30]

Figure 9 (panel a) reveals that male LFP rates drop once their spouse reaches the age of 70 and thus also becomes eligible. The drop in male LFP rates did not exist before the pension was introduced. Confirming the graphical evidence, RD estimates in table 8 (column 1) report a significant LFP drop of 3.3 percentage points among eligible married men whenever their spouse becomes eligible. Table 8 also shows that these results survive several placebo tests and are robust against varying bandwidths.

To make sure that the LFP drop is not driven by the fact that own age correlates with spousal age, we also present results for own age-adjusted LFP rates. Similar to Garcia-Miralles and Leganza (2021), we construct own age-adjusted LFP rates by running a regression of LFP on own age and taking the residual.[31] The results depicted in figure 9 (panel b) and corresponding estimates in table 8 (column 2) show that the estimated drop in age-adjusted LFP rates closely resembles the non-age-adjusted LFP drop.

We conclude that these results are strongly suggestive for within-family spillovers and are in line with previous findings on joint retirement decisions of couples. Our finding is further

[30]Figure B.5 in the appendix shows the average age of married men (panel a) and the sample size (panel b) conditional on spousal age. Give the construction of our sample, the average age of married men is always above 70 irrespective of spousal age. Thus, age differences within couples are largest for the lower end of spousal ages (panel a). The larger share of observations around the age 70 threshold (panel b) indicates that lower age differences among couples are more frequent.

[31]We run the regression for all men aged 50 to 90 once for the UK census in 1911 and separately for 1901. Age in years enters the regression as a range of dummy variables.

corroborated by anecdotal historical evidence suggesting that the OPA was particularly beneficial for married couples with two pensions (say 2 x 5 *shillings*) compared to otherwise identical single households receiving one pension (5 *shillings*) but with less potential for synergies (see Pugh, 2002, p. 790).

4.3 Validity and Robustness of the RD

Several sensitivity checks and falsification tests support the validity of our RD design. Based on these exercises, the estimates presented above can be given a causal interpretation.

4.3.1 Smoothness Analysis

We start verifying the validity of the RD by testing whether observable characteristics other than the introduction of the OPA can explain the discontinuity in LFP. Therefore, we test whether pre-determined covariates exhibit a discontinuity at the threshold. Our smoothness analysis includes the share of individuals living in urban areas, the share of foreign born, the share of disabled persons, the number of own children in the household, the share of married or individuals living alone (similar to the summary statistics reported in table 1). Two variables (share urban, number of children) evolve smoothly around the age-based eligibility threshold. Four variables, however, show statistically significant discontinuities (table 9 and figure B.6). Predominantly, these discontinuities are either negligibly small (share disabled, share single households) or not robust (share married).[32]

The only discontinuity that seems relevant in terms of magnitude is the increase in the share of foreign born people by 1.3 percentage points. However, we are confident that this discontinuity does not threaten the validity of the RD for two reasons. First and most importantly, we see significant and similar-sized LFP declines for natives as well as for foreign born people (table A.1 and figure B.7). Thus, the LFP decline is independent of the nativity status. Second, we observe exactly the same increase in the share of foreign born (also by 1.3 percentage points) at age 70 in 1901 before the pension was introduced. Therefore, the foreign born discontinuity

[32]The discontinuity for the share of married people disappears once we include quadratic age polynomials.

does not seem to be driven by the availability of the pension but can be explained by age-heaping (round number bunching). Figure B.8 shows that the number of foreign born people exhibits a very pronounced age-heaping pattern in 1901 as well as in 1911, which directly translates into an increase in the share of foreign born at round ages given that age-heaping is less pronounced for natives.

4.3.2 Placebo Tests

Next, we conduct two placebo tests to show that LFP effects do not appear at any arbitrarily chosen age cutoff. First, we run the baseline RD specification on the 1901 census to ensure that the abrupt LFP decline measured after the introduction of the OPA in 1911 does not occur in 1901. Table 10 (upper panel) indicates that there is no sizable LFP decline in 1901, consistent with graphical evidence in figure 2 that indicates a smooth LFP decline over age before the pension was introduced. Table 10 also shows that there is no substantial drop in LFP at age 60 in the 1901 census. Since the birth cohort that reached age 60 in 1901 eventually reached age 70 in 1911, this robustness check rules out that individuals who play a key role for the identifying variation in the RD design already exhibited an LFP decline at earlier ages before the OPA was introduced.

Second, we repeat the analysis for arbitrarily chosen placebo age cutoffs in 1911. RD estimates in table 10 (lower panel) show that we do not observe a similar decline in LFP rates at any hypothetical age cutoff other than the true eligibility threshold at age 70.[33]

4.3.3 Manipulation of the Running Variable

Even though we cannot formally[34] rule out that age manipulation affects our results, we consider it to be a negligible threat to identification for three reasons. First, pensions were not

[33]This result holds for any other hypothetical age cutoff. Additional placebo tests on arbitrarily chosen age cutoffs are available from the authors upon request.

[34]We have conducted the formal manipulation test proposed by Cattaneo et al. (2018). Based on this test, we have to reject the null hypothesis of no difference in the density at age 70. However, we consider this result to have little informational value in our context, since we have to reject the null hypothesis for all age-groups in our sample except for three ages (86, 77 and 67).

granted based on the age provided in the census, but on the information from official birth certificates that existed in England and Wales for more than seventy years (since at least 1837) when the pension was implemented. People in England and Wales had therefore little incentive to manipulate their age in the census in order to become eligible. Second, although there is some evidence of age-heaping[35] at the eligibility age in 1911, it is less pronounced compared to the US at the time or in England and Wales ten years earlier (figure B.9). Thus, there is no evidence that the introduction of old-age assistance resulted in willful age-manipulation. Third, we have conducted Donut-RD regressions to test how much our results are driven by observations in close proximity to the cut-off. If present, age-misreporting should arguably be most pronounced for age-groups closely below or above the age cut-off. Table 11 shows that the results remain robust if we exclude these observations.

4.3.4 Anticipatory Effects

Another potential challenge for the age-RD is anticipatory behavior. People may respond to the pension by reducing their labor supply even before the eligibility age (e.g. due to income effects), which could bias our estimates. Neither figure 2, nor Donut-RD estimates in table 11 nor the DiD representation in figure B.3 suggest that anticipation effects are a relevant concern. To test the role of anticipatory effects in even more detail, we run an extended DiD model that compares LFP changes between 1901 and 1911 for all ages (50 to 90) to the reference age of 69 (just below the eligibility age). Under the assumption that work becomes more burdensome with age, we would expect that most anticipatory retirement (= retirement before the eligibility age) would occur in the age groups slightly below the newly introduced eligibility cutoff.

Figure 10 reveals that there is little evidence for anticipatory retirement before age 70. In fact, the LFP rates of people from 50 to 69 have moved in tandem between 1901 and 1911. In contrast, LFP rates of people 70+ have clearly decreased relative to LFP rates of people at age 69. These differences fade out in very old age (80+) because LFP rates in this age range

[35] Age-heaping itself does not result in substantial LFP discontinuities, e.g. there is no significant LFP discontinuity at age 60 in England and Wales in 1911 despite the substantial age-heaping.

were relatively low even before the pension was introduced, leaving little room for additional retirement.

4.3.5 Bandwidth Choice, Polynomials, Estimation Technique and Inference

We finally present several adjustments to the baseline specification that have been proposed in the RD literature to verify the robustness of the estimates and the validity of statistical inference (table 12). First, the estimates are not sensitive to bandwidth choice and only change little when making the bandwidth arbitrarily large. Second, changing the polynomial degree when modeling LFP as a function of age does not indicate sizable changes in the coefficients. Third, parametric baseline estimates are also robust to local non-parametric estimation techniques (conventional and bias-corrected) as suggested by Calonico et al. (2014a,b). Fourth, instead of clustering by the (discrete) running variable, as has been proposed by Lee and Card (2008) and is common practice in many empirical applications (e.g. Card et al., 2008, 2009), we also report standard errors based on the BME (bounded misspecification error) procedure developed by Kolesár and Rothe (2018). The BME standard errors are slightly larger, throughout, but do not affect our conclusions. Overall, changes in estimates (and standard errors) are moderate even when undertaking considerable changes in bandwidth, age polynomials, weighting scheme (uniform vs. triangular kernel) or estimation procedure.

5 Discussion and Implications

5.1 Mechanisms

In this section, we argue that the estimated labor supply response to the OPA is consistent with substitution effects and constrained income effects. We cannot rule out that newly established social norms also played a role but consider it to be less likely.

The OPA includes an indirect tax (substitution effect) and a transfer component (income effect). The clearest evidence for the relevance of the substitution effect is the clear and abrupt decline in LFP at the eligibility threshold. At this age, the OPA creates a kink in the lifetime budget constraint that is driven by the existence of the means test. The means test puts an implicit tax on labor earnings, which makes labor (relative to leisure) less attractive, and results in additional retirement starting at age 70. The relevance of substitution effects for labor supply decisions in a setting with earnings-tested retirement benefits has recently been documented in the US context (Gelber et al., 2021a,b).

The relevance of the transfer component (income effect) of the OPA is less clear. In a model with perfect foresight, the increase in the lifetime budget would also affect labor supply at younger ages. However, the results of section 4.3.4 suggest that labor supply below the eligibility age did not respond to the OPA. One compelling reason for the absence of labor supply responses before the eligibility age are liquidity constraints. If people lack assets to cover the period between their preferred labor market exit and pension eligibility –and are not able to borrow against their future pension benefits– they can only leave the labor market once they reach the eligibility age. The target population of the OPA was probably at least partly liquidity constrained even before reaching the eligibility age. Moreover, at the time, credit markets were underdeveloped compared to today's standards.[36] Further potential reasons for the lack of LFP responses before reaching the eligibility age are myopic behavior or uncertainty about future changes to pension legislation. Given these constraints, any income effect would also show up as a decline in LFP rates at the eligibility age making it impossible to separate substitution from constrained income effects.

The LFP decline at the eligibility threshold could also be consistent with a newly created retirement norm but we consider this explanation to be less compelling. Recent studies have shown that retirement behavior in long-standing pension systems is not only driven by financial incentives, but also by reference points (see e.g. Behaghel and Blau, 2012; Seibold, 2021). The OPA was a salient reform with the potential to establish a new retirement norm, although perhaps the two years that lie in between the introduction of the OPA (1909) and the documented labor supply response (census 1911) are too short to create a new reference point. More importantly, the finding that people at the bottom of the occupational earnings distribution respond more strongly to the OPA (section 4.2.1) speaks against norms as a plausible explanation. If

[36] Accoding to the Jordà-Schularick-Taylor Macrohistory Database (Jordà et al., 2016), the ratio of household loans to GDP in the UK was 3.5% in 1911, compared to 68.0% in 2017.

norms were driving the result, we would expect that the response would be more homogeneous across earnings groups, or even more pronounced at the top given the finding by Behaghel and Blau (2012) that people with higher cognitive skills respond more strongly to newly created reference points.

5.2 Welfare Implications

A priori, the welfare effect of the OPA is unclear. On the one hand, the introduction of old-age assistance lowered societal welfare by distorting labor supply decisions (efficiency loss). On the other hand, redistribution from rich tax payers to poor elderly people increased societal welfare if the society valued equality (equity gain).

The efficiency loss from the pension was low for two reasons.[37] First, the labor supply response was modest relative to the size of the program. Assuming that, in absence of a pension, LFP of people 70+ would have evolved as for 65-69 year old's, less than 10% of pension beneficiaries retired as a response to the OPA. Second, only high-income people paid taxes (around 1 million out of 36 million inhabitants) and for those taxes rates were low (max. 8.3%). Thus, fiscal externalities of the OPA were limited. The lower labor supply of poor elderly people had arguably little effect on government revenues and substantial behavioral responses of top-income earners to increased taxation are unlikely given the low tax rates.

In contrast, the equality gain was immense. In a time of high income inequality, the OPA generated substantial transfers from wealthy tax payers (annual income: > £5000) to poor elderly pension recipients (annual income before the pension: < £30) (see section 2). Therefore, we conclude that the OPA was welfare enhancing even if the society had only a small preference for equality.

[37] See appendix D, for an in-depth welfare analysis, including details for the underlying calculation, based on the methodology pioneered by Hendren (2016) and Hendren and Sprung-Keyser (2020)

5.3 Effect Size and Policy Implications

We now conduct a simple back-of-the-envelope calculation to put the effect size of our LFP estimates into today's context. Specifically, we quantify by how much a reversed OPA-type intervention would offset the demographic pressure on the current pension system in the UK. For this purpose, we assume a hypothetical policy that raises the LFP rate by 6 percentage points (the reverse effect of the OPA) in the relevant age range and then examine how this affects the ratio of retirees relative to the number of people in the labor force (retiree-to-LFP ratio) until 2030 and 2050. Since the external validity of our estimates is limited, we do not specify how such a hypothetical policy would need to look like today in order to raise LFP rates accordingly.

The current (2019) retiree-to-LFP ratio in the UK is 31.7, implying that there are about 32 retirees aged 65 and over for every 100 persons in the labor force.[38] Holding LFP rates fixed, the retiree-to-LFP ratio will rise to 38 until 2030 based on the medium fertility variant of the UN population projections (United Nations, 2019). To keep the ratio at its 2019 level (31.7), it would take almost 2.2 million people who switch their status from retiree (numerator) to labor force participant (denominator) in 2030.

We assume that a reversed OPA-type policy would mostly affect the LFP rate of people aged 65 – 74.[39] This age group is predicted to consist of 7.8 million people in 2030 and raising their LFP rate by the (reverse) OPA effect of 6 percentage points implies an expansion of the labor force by around 0.46 million older workers. This is less than one quarter (0.46 out of 2.2 million) of the size required to keep the retiree-to-LFP ratio constant.

Looking further ahead until 2050, a reversed OPA policy targeting the age group 65 – 74 would only compensate one eighth of the increase needed to keep UK's retiree-to-LFP ratio

[38]We calculate this ratio by dividing the number of people outside the labor force and aged 65+ (retirees) over the number of of people inside the labor force at all ages. To obtain this measure we combine statistics on LFP rates (OECD, 2021) with population statistics (United Nations, 2019). The size of the retiree-to-LFP ratio is similar to the demographic "old-age to working-age" ratio (31.7 in 2019) that is defined by the number of people 65+ relative to people in working age 20 to 64 (OECD, 2019, p. 176).

[39]The fact that these people are eligible for a State Pension– a successor program of the OPA– likely contributes to the currently low LFP rates in this age group (LFP rate of 17% in 2019). In line with our OPA extended DiD estimates (figure 10), we expect that changes to the State Pension, such as the already planned increase in the eligibility age from 65 to 68 until the mid 2040s, will mostly affect the labor supply of people that are in the age range close to the current eligibility age ($\approx$ 65 years in 2019).

constant.[40] Our back-of-the-envelope calculation demonstrates that future reforms to stabilize the retiree-to-LFP ratio have to induce much larger labor supply responses than the OPA.

6 Conclusions

In this paper, we take advantage of a unique historical reform, the Old Age Pension Act of 1908 in the UK, to study how the introduction of means-tested old-age assistance affects labor supply. Using recently digitized full-count census data, we estimate the causal impact of the policy on the labor supply of older workers by exploiting variation around the newly created age-based eligibility threshold.

We find a considerable reduction of labor supply in the local environment of the eligibility threshold after the pension was implemented. When reaching the eligibility age of 70, the labor force participation rate declines by 6 percentage points or 13%. This decline is strongly driven by older workers who directly retire from work (and not unemployment). Labor supply reductions are larger when occupational earnings are low and when family ties, as a proxy for private transfers, are weak. Our results are also suggestive for joint retirement decisions of couples.

The transparent redistributive design of the reform allows us to study the overall welfare implications of the program. We argue that, despite the considerable labor supply decline, the introduction of old-age assistance in the UK was welfare enhancing if the society was at least slightly averse to inequality. The pension and its financing via higher top income tax rates created only limited behavioral responses, relative to the immense size of the program. In contrast, equity gains were large as the policy redistributed from the top to the bottom of the earnings distribution in a time of relatively high income inequality and only limited social security.

[40]In 2050, UK's retiree-to-LFP ratio will have risen to 46 (using 2019 LFP rates) and it would then require almost 4 million people who remain in the labor force instead of retiring to keep the ratio constant. The age group 65 – 74 is predicted to have grown to 8.2 million people, implying that the reverse OPA policy would increase the labor force by 0.49 million people. This compensates one eighth (0.49 out of 4 million) of the increase required to keep the ratio constant.

The historical setting allows us to credibly identify the labor supply effect of an old-age assistance program and thus extends a large body of literature that quantifies the incentive effects of marginal changes in existing pension schemes. Our results are policy-relevant, not only for developing countries without universal coverage by old-age assistance (e.g. in Sub-Sahara Africa), many of which are also characterized by high income inequality and limited public welfare spending. Placing the magnitude of the labor supply response of the OPA into the context of today's well-established pension schemes also suggests that future pension reforms that aim to stabilize the retiree-to-worker ratio have to be even more profound than the OPA. Raising the retirement age by perhaps one or two years, thus, will hardly be enough to meet the challenges posed by population aging.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy 122*(3), 467–506.

Anderson, M., C. Dobkin, and T. Gross (2012). The effect of health insurance coverage on the use of medical services. *American Economic Journal: Economic Policy 4*(1), 1–27.

Anderson, M. L., C. Dobkin, and T. Gross (2014). The effect of health insurance on emergency department visits: Evidence from an age-based eligibility threshold. *Review of Economics and Statistics 96*(1), 189–195.

Arthi, V., B. Beach, and W. W. Hanlon (2021). Recessions, mortality, and migration bias: Evidence from the lancashire cotton famine. *American Economic Journal: Applied Economics* (forthcoming).

Atalay, K. and G. F. Barrett (2015). The impact of age pension eligibility age on retirement and program dependence: Evidence from an australian experiment. *The Review of Economics and Statistics 97*(1), 71–87.

Atalay, K., G. F. Barrett, and P. Siminski (2019). Pension incentives and the joint retirement of couples: Evidence from two natural experiments. *Journal of Population Economics 32*(3), 735–767.

Atkinson, A. B. (2005). Top incomes in the UK over the 20th century. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 168*(2), 325–343.

Atkinson, A. B. (2007). The distribution of top incomes in the united kingdom 1908-2000 (chapter 4). In A. B. Atkinson and T. Piketty (Eds.), *Top incomes over the twentieth century. A contrast between continental european and english-speaking countries*, pp. 82 – 140. Oxford University Press.

Baker, M. (2002). The retirement behavior of married couples: Evidence from the spouse's allowance. *The Journal of Human Resources 37*(1), 1–34.

Battistin, E., A. Brugiavini, E. Rettore, and G. Weber (2009). The retirement consumption puzzle: Evidence from a regression discontinuity approach. *American Economic Review 99*(5), 2209–26.

Beach, B. and W. W. Hanlon (2019). Censorhip, familiy planning, and the historical fertility transition. *NBER Working Paper No. 25752*.

Behaghel, L. and D. M. Blau (2012). Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age. *American Economic Journal: Economic Policy 4*(4), 41–67.

Blau, D. M. and R. T. Riphahn (1999). Labor force transitions of older married couples in germany. *Labour Economics 6*(2), 229–252.

Boldrin, M. and L. E. Jones (2002). Mortality, fertility, and saving in a malthusian economy. *Review of Economic Dynamics 5*(4), 775–814.

Börsch-Supan, A. (2000). Incentive effects of social security on labor force participation: Evidence in germany and across europe. *Journal of Public Economics 78*(1-2), 25–49.

Boyer, G. R. (2016). "Work for their prime, the workhouse for their age": Old Age Pauperism in Victorian England. *Social Science History 40*(1), 3–32. Publisher: Cambridge University Press.

Brown, K. M. (2013). The link between pensions and retirement timing: Lessons from california teachers. *Journal of Public Economics 98*, 1 – 14.

Bureau of Labor Statistics (1929). Status of Old-Age Pension Legislation in the United States. *Monthly Labor Review 29*(1), 21–28. Publisher: Bureau of Labor Statistics, U.S. Department of Labor.

Cain, M. (1983). Fertility as an Adjustment to Risk. *Population and Development Review 9*(4), 688–702.

Caldwell, J. C. (1982). *Theory of fertility decline*. London: Academic Press.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014a). Robust Data-Driven Inference in the Regression-Discontinuity Design. *Stata Journal 14*(4), 909–946.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014b). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Card, D., C. Dobkin, and N. Maestas (2008). The impact of nearly universal insurance coverage on health care utilization: Evidence from medicare. *American Economic Review 98*(5), 2242–58.

Card, D., C. Dobkin, and N. Maestas (2009). Does medicare save lives? *The Quarterly Journal of Economics 124*(2), 597 – 636.

Carpenter, C. and C. Dobkin (2009). The effect of alcohol consumption on mortality: Regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics 1*(1), 164–182.

Carpenter, C. and C. Dobkin (2015). The minimum legal drinking age and crime. *The Review of Economics and Statistics 97*(2), 521–524.

Carpenter, C. and C. Dobkin (2017). The minimum legal drinking age and morbidity in the united states. *The Review of Economics and Statistics 99*(1), 95–104.

Casson, W. A. (1908). *Old-age pensions act, 1908: Together with the text of the regulations made thereunder dated 15th october, 1908, and official circulars and instructions for the guidance of pension authorities by the local government boards of england, scotland, and ireland; annotated and explained, with historical introduction*. C. Knight & Company, Limited.

Cattaneo, M. D., M. Jansson, and X. Ma (2018, March). Manipulation Testing Based on Density Discontinuity. *The Stata Journal 18*(1), 234–261. Publisher: SAGE Publications.

Chamberlain, J. P. and S. Pierson (1924). CURRENT LEGISLATION: OLD AGE PENSION LEGISLATION. *American Bar Association Journal 10*(2), 109–111. Publisher: American Bar Association.

Costa, D. L. (1998). The evolution of retirement. In *The evolution of retirement: An american economic history, 1880-1990*, pp. 6–31. NBER and University of Chicago Press.

Danzer, A. M. (2013). Benefit generosity and the income effect on labour supply: Quasi-experimental evidence. *The Economic Journal 123*(571), 1059–1084.

Department of Labour Statistics (1913). Sixteenth Abstract of Labour Statistics of the United Kingdom. Abstract of Labour Statistics.

Department of Labour Statistics (1915). Seventeenth Abstract of Labour Statistics of the United Kingdom. Abstract of Labour Statistics.

Diamond, P. and E. Saez (2011). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives 25*(4), 165–90.

Feinstein, C. (1990). New Estimates of Average Earnings in the United Kingdom, 1880-1913. *The Economic History Review 43*(4), 595–632.

Fetter, D. K. and L. M. Lockwood (2018). Government old-age support and labor supply: Evidence from the old age assistance program. *American Economic Review 108*(8), 2174–2211.

Finkelstein, A. (2019). Welfare analysis meets causal inference: A suggested interpretation of hendren. *Note, MIT Economics*.

Fitzpatrick, M. D. and T. J. Moore (2018). The mortality effects of retirement: Evidence from social security eligibility at age 62. *Journal of Public Economics 157*, 121 – 137.

Friedberg, L. (1999). The effect of old age assistance on retirement. *Journal of Public Economics 71*(2), 213 – 232.

Garcia-Miralles, E. and J. Leganza (2021). Joint Retirement of Couples: Evidence from Discontinuities in Denmark. *Mimeo*.

Gelber, A. M., D. Jones, D. W. Sacks, and J. Song (2021a). Using Non-linear Budget Sets to Estimate Extensive Margin Responses: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics* (forthcoming).

Gelber, A. M., D. Jones, D. W. Sacks, and J. Song (2021b). The Employment Effects of the Social Security Earnings Test. *Journal of Human Resources* (forthcoming).

Gustman, A. L. and T. L. Steinmeier (2000). Retirement in dual-career families: A structural model. *Journal of Labor Economics 18*(3), 503 – 545.

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy 30*(1), 51–89.

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics 187*.

Hendren, N. and B. Sprung-Keyser (2020). A Unified Welfare Analysis of Government Policies*. *The Quarterly Journal of Economics 135*(3), 1209–1318.

Higgs, E., C. Jones, K. Schürer, and A. Wilkinson (2013). Integrated census microdata (i-cem) guide; 1851-1911. *University of Essex, Department of History*.

House of Commons (1911). Revenue and expenditure (England, Scotland, and Ireland). *House of Commons Papers 220*(45), 16.

Human Mortality Database (2018). University of california, berkeley and max planck institute for demographic research, rostock. *available at www.mortality.org or www.humanmortality.de*(data accessed on October 1, 2018).

Johnson, P. (1994). The employment and retirement of older men in england and wales, 1881 – 1981. *The Economic History Review 47*(1), 106–128.

Jordà, Ò., M. Schularick, and A. M. Taylor (2016, May). Macrofinancial History and the New Business Cycle Facts. *NBER*, 213–263.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Kolesár, M. and C. Rothe (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review 108*(8), 2277–2304.

Krueger, A. B. and J.-S. Pischke (1992). The effect of social security on labor supply: A cohort analysis of the notch generation. *Journal of Labor Economics 10*(4), 412–37.

Lee, D. S. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

Leibenstein, H. (1957). *Economic backwardness and economic growth : Studies in the theory of economic development*. John Wiley and Sons.

Liebman, J. B., E. F. P. Luttmer, and D. G. Seif (2009). Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities. *Journal of Public Economics 93*(11 - 12), 1208 – 1223.

Manoli, D. and A. Weber (2016). Nonparametric evidence on the effects of financial incentives on retirement decisions. *American Economic Journal: Economic Policy 8*(4), 160–82.

Mastrobuoni, G. (2009). Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities. *Journal of Public Economics 93*(11 - 12), 1224 – 1233.

Minnesota Population Center (2018). Integrated Public Use Microdata Series, International: Version 7.0.

Murray, B. (2009). The Peoples Budget a century on. *Journal of Liberal History Liberal Democrat History Group*(64), 4 – 13.

OECD (2019). Pensions at a Glance 2019: OECD and G20 Indicators. Technical report, OECD Publishing, Paris.

OECD (2021). Labour Market Statistics: Labour Force Statistics by Sex and Age: Indicators. *https://www.oecd-ilibrary.org/content/data/data-00310-en* (data accessed on May 12, 2021).

Old Age Pensions Committee (1919). Old age pensions: Appendix to the report of the departmental committee on old age pensions including minutes of evidence. Technical report, London.

Parsons, D. O. (1991). Male Retirement Behavior in the United States, 1930–1950. *The Journal of Economic History 51*(3), 657–674. Publisher: Cambridge University Press.

Pugh, M. (2002). Working-Class Experience and State Social Welfare, 1908–1914: Old Age Pensions Reconsidered. *The Historical Journal 45*(4), 775–796. Publisher: Cambridge University Press.

Seibold, A. (2021). Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities. *American Economic Review 111*(4), 1126–65.

Spender, J. A. (1892). The state and pensions in old age. *London, SWAN SONNENSCHEIN & CO*.

Thane, P. (2000). *Old age in english history: Past experiences, present issues* (1 edition ed.). Oxford UK ; New York: Oxford University Press.

The Times (1909). Old-age pensions. *The Times* (London), page 9.

United Nations (2019). World population prospects 2019, online edition (rev. 1). *available at https://population.un.org*(data downloaded on May 10, 2021).

Appendix

Tables

Table 1: Summary Statistics by Age (Census 1911)

	69		70	
	Mean	S.D.	Mean	S.D.
Share Female	0.55	0.50	0.56	0.50
Share Urban	0.10	0.30	0.10	0.30
Share Married	0.50	0.50	0.47	0.50
Share Foreign Born	0.04	0.20	0.06	0.24
Share Disabled	0.02	0.13	0.02	0.14
N Children in the Household	0.8	1.1	0.7	1.0
Share Single Households	0.07	0.26	0.08	0.28
Labor Force Participation Rate	0.46	0.50	0.39	0.49
Observations	150,293		140,288	

	65 - 69		70 - 74	
	Mean	S.D.	Mean	S.D.
Share Female	0.55	0.50	0.57	0.50
Share Urban	0.10	0.30	0.10	0.29
Share Married	0.54	0.50	0.43	0.50
Share Foreign Born	0.04	0.20	0.05	0.22
Share Disabled	0.02	0.13	0.02	0.14
N Children in the Household	0.9	1.2	0.7	1.0
Share Single Households	0.07	0.25	0.09	0.29
Labor Force Participation Rate	0.48	0.50	0.34	0.47
Observations	803,208		551,100	

Source: UK Census wave 1911 and IPUMS. *Note:* Reported values for men and women. The upper panel reports values marginally below and above the age cutoff (69 and 70). The lower panel reports values for the baseline estimation sample with 5 age-years below the cutoff (65 - 69) and above the cutoff (70 - 74).

Table 2: Occupational Composition (Census 1911)

Occupation	Observations	Percent
Legislators and Managers	10,781	0.8
Professionals	17,262	1.3
Technicians	9,292	0.7
Clerks	14,500	1.1
Service Workers	151,528	11.2
Agriculture and Fishery	105,472	7.8
Crafts and Related Trades	170,959	12.6
Machine Operators and Assemblers	39,143	2.9
Elementary Occupations	54,348	4.0
Armed Forces	1,881	0.1
Active	575,166	42.5
Inactive	779,142	57.5
Total	1,354,308	100

Source: UK Census wave 1911 and IPUMS. *Note:* Reported values on occupations based on ISCO classification at the 1-digit level for individuals aged 65 - 74.

Table 3: RD Estimates of Labor Force Participation at the Age Cutoff (Baseline)

	Baseline RD Estimate	Relative Decline
	-0.060** (0.006)	-13% (-0.060/0.456)
Observations	1,354,308	

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimate of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Linear polynomials in age. Specification uses a bandwidth of 5 age-years to the left (age 65-69, N: 803,208) and to the right (age 70-74, N: 551,100) of the age cutoff and uniform weighting on all observations. Relative Decline is measured by dividing the RD estimate by the mean of the LFP rate at age 69.**, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table 4: Difference-in-Differences Estimates of Labor Force Participation

	(1)	(2)	(3)
DiD Estimate	-0.045** (0.003)	-0.040** (0.002)	-0.040** (0.002)
Controls	–	✓	✓
Controls X Census Year	–	–	✓
Observations	715,714	698,921	698,921

Source: UK Census wave 1891, 1901, 1911 and IPUMS. *Note:* Estimates for men and women aged 69 and 70. **, * denotes significance at the 1% and 5% level respectively. Standard errors in parentheses. The dependent variable is the labor force participation rate.

Table 5: RD Estimates of Labor Supply Response

	Working	Unemployed
	(1)	(2)
RD Coefficient	-0.057** (0.006)	-0.002* (0.001)
Relative Decline	-13% (-0.057/0.430)	-8% (-0.002/0.025)
Observations	1,354,308	

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Estimates based on separate definitions of labor force status: LFP = 1 for individuals who are working (column 1) or LFP = 1 for individuals who are unemployed/formerly employed (column 2). All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. Relative Decline is measured by dividing the RD estimate by the mean of the LFP rate at age 69. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table 6: RD Estimates Stratified by Family Background

	RD Coefficient	Relative Decline	Observations
	Children		
Children in Household	-0.051** (0.006)	-12% (-0.051/0.417)	627,733
No Children in Household	-0.069** (0.007)	-14% (-0.069/0.489)	726,575
z-Statistic (Significant Differences)	2.0*		
	Marital Status		
Married	-0.056** (0.006)	-12% (-0.056/0.478)	679,145
Non-Married	-0.062** (0.007)	-14% (-0.056/0.433)	675,163
z-Statistic (Significant Differences)	0.6		
	Household Size		
Single Person Households	-0.095** (0.007)	-20% (-0.095/0.475)	102,938
Multiple Person Households	-0.057** (0.006)	-13% (-0.057/0.454)	1,251,370
z-Statistic (Significant Differences)	-3.9**		

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Estimates for sub-samples along the lines of three types of family background variables. All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. Relative Decline is measured by dividing the RD estimate by the mean of the LFP rate at age 69. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years. Testing for differences between two coefficients with known variances yields a z-test statistic (assumed to be normally distributed), where β is the estimated coefficient for the respective sub-sample and *SE* denotes the corresponding standard errors. This yields $z = \frac{\beta_C - \beta_{NC}}{\sqrt{SE^2_{\beta_C} + SE^2_{\beta_{NC}}}} = 2.0$ for testing between children/no children, $z = \frac{\beta_M - \beta_{NM}}{\sqrt{SE^2_{\beta_M} + SE^2_{\beta_{NM}}}} = 0.6$ for testing between married/non-married, and $z = \frac{\beta_{SP} - \beta_{MP}}{\sqrt{SE^2_{\beta_{SP}} + SE^2_{\beta_{MP}}}} = -3.9$ for testing between single/multiple person households.

Table 7: RD Estimates by Gender

	RD Coefficient	Relative Decline	Observations
Men	-0.079** (0.011)	-10% (-0.079/0.765)	602,458
Women	-0.032** (0.004)	-16% (-0.032/0.199)	751,850
z-Statistic (Significant Differences)	-4.0**		

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Estimates for the sub-samples of men (column 1) and women (column 2). All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. Relative Decline is measured by dividing the RD estimate by the mean of the LFP rate at age 69. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years. Testing for differences between two coefficients with known variances yields a z-test statistic (assumed to be normally distributed), where β is the estimated coefficient for the respective sub-sample and SE denotes the corresponding standard errors. This yields $z = \frac{\beta_M - \beta_W}{\sqrt{SE^2_{\beta_M} + SE^2_{\beta_W}}} = -4.0$ for testing between Men/Women.

Table 8: RD Estimates for Married Men (Age $\geq$ 70) by Spousal Eligibility

	(1) LFP	(2) Age-Adjusted LFP
True Age Cutoff: Age of Spouse $\geq$ 70, Census 1911		
Bandwidth: 5 Age-Years (Baseline)	-0.033** (0.006)	-0.033** (0.006)
Observations	90,815	
Bandwidth: 20 Age-Years (quadratic)	-0.036** (0.004)	-0.040** (0.004)
Observations	161,008	
Bandwidth: 10 Age-Years (quadratic)	-0.034** (0.007)	-0.033** (0.007)
Observations	136,478	
Bandwidth: 5 Age-Years (quadratic)	-0.023* (0.010)	-0.022 * (0.010)
Observations	90,815	
Bandwidth: 4 Age-Years (linear)	-0.033** (0.006)	-0.032** (0.006)
Observations	76,159	
Bandwidth: 3 Age-Years (linear)	-0.031** (0.009)	-0.029** (0.008)
Observations	59,619	
Placebo Cutoffs		
Age of Spouse $\geq$ 60, Census 1911	-0.014 (0.017)	-0.017 (0.017)
Observations	28,085	
Age of Spouse $\geq$ 65, Census 1911	0.015 (0.010)	0.013 (0.008)
Observations	57,616	
Age of Spouse $\geq$ 60, Census 1901	0.018(0.011)	0.017 (0.010)
Observations	26,894	
Age of Spouse $\geq$ 70, Census 1901	0.012 (0.008)	0.013 (0.07)
Observations	73,370	

Source: UK Census wave 1901 and 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate of married men (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70) of the respective spouse. The sample is restricted on men having reached the eligibility age of 70 (71+). Estimates are computed separately for LFP (1) and LFP adjusted by own age (2). All regressions (lower panel) use a bandwidth of 5 age-years to the left and to the right of the cutoff, a linear polynomial in age, and uniform weighting on all observations. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table 9: Smoothness Analysis

Share Urban	0.003 (0.002)
Share Foreign Born	0.013** (0.004)
Share Disabled	0.001* (0.000)
N Children	-0.007 (0.004)
Share Married	-0.012* (0.005)
Share Single Households	0.005** (0.001)
Observations	1,354,308

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the respective observable (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table 10: Placebo Tests

Age Cutoff	RD Coefficient	Observations
	Census 1901	
70	-0.009* (0.004)	1,041,111
60	0.001 (0.003)	1,889,492
	Census 1911	
69	-0.017 (0.012)	1,446,757
65	-0.004 (0.002)	1,820,152
60	0.001 (0.002)	2,293,676

Source: UK Census wave 1901 and 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for different placebo age cutoffs (as indicated). All regressions use a bandwidth of 5 age-years to the left and to the right of the cutoff, a linear polynomial in age, and uniform weighting on all observations. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table 11: Donut-RD Estimates

Excluded Ages	RD Coefficient	Observations
1-year (69, 71)	-0.077** (0.004)	1,063,727
2-years (68, 69, 71, 72)	-0.076** (0.003)	799,849
3-years (67, 68, 69, 71, 72, 73)	-0.087** (0.000)	534,770

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

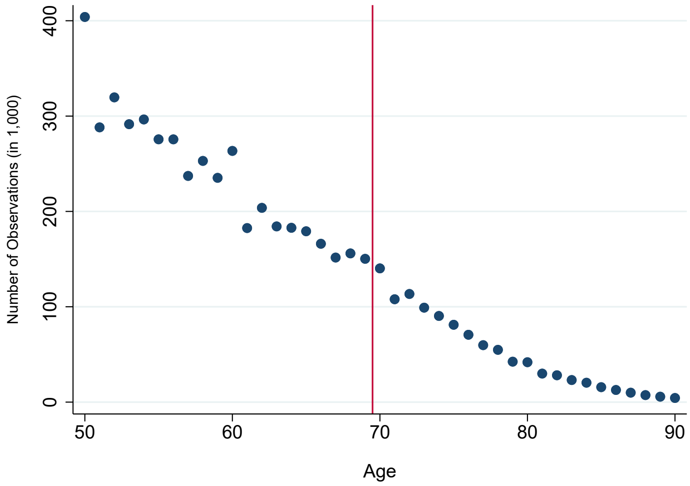
Table 12: Alternative Specifications and Inference for the Baseline RD

	Parametric				Local Non-Parametric	
	Clustered Standard Errors		BME Standard Errors		Clustered Standard Errors	
	Linear	Quadratic	Linear	Quadratic	Conventional	Bias-Corrected
	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth (Age-Years)						
2	-0.048** (0.000)	–	-0.048** (0.003)	–	–	–
3	-0.056** (0.004)	-0.038** (0.000)	-0.056** (0.007)	-0.038** (0.006)	–	–
4	-0.058** (0.005)	-0.053** (0.005)	-0.058** (0.008)	-0.053** (0.006)	-0.054** (0.003)	-0.038** (0.001)
5	-0.060** (0.006)	-0.053** (0.003)	-0.060** (0.008)	-0.053** (0.007)	-0.056** (0.004)	-0.050** (0.005)
6	-0.063** (0.007)	-0.055** (0.003)	-0.063** (0.009)	-0.055** (0.007)	-0.058** (0.005)	-0.052** (0.003)
7	-0.064** (0.008)	-0.056** (0.004)	-0.064** (0.010)	-0.056** (0.007)	-0.059** (0.006)	-0.053** (0.003)
8	-0.066** (0.008)	-0.058** (0.004)	-0.066** (0.011)	-0.058** (0.007)	-0.061** (0.007)	-0.055** (0.003)
9	-0.070** (0.009)	-0.055** (0.005)	-0.070** (0.013)	-0.055** (0.008)	-0.062** (0.007)	-0.056** (0.003)
10	-0.072** (0.009)	-0.057** (0.005)	-0.072** (0.013)	-0.057** (0.009)	-0.064** (0.008)	-0.054** (0.003)
20	-0.110** (0.012)	-0.061** (0.007)	-0.110 (0.057)	-0.061** (0.012)	-0.090** (0.011)	-0.055** (0.007)

Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Parametric estimates in column (1) – (4) use uniform weighting of all observations. Local non-parametric estimates in column (5) – (6) use triangular kernel weighting that puts more weight on observations closer to the cutoff. Estimates in column (6) are based on the bias correction proposed by Calonico et al. (2014a,b). Standard errors are clustered at age in years in column (1) – (2) and column (5) – (6), while bounded mis-specification standard errors (BME) proposed by Kolesár and Rothe (2018) are reported in column (3) – (4). **, * denotes significance at the 1% and 5% level respectively.

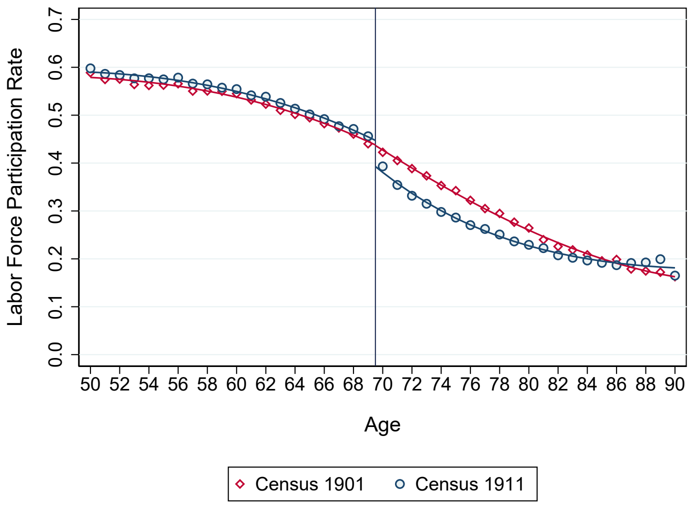
Figures

Figure 1: Number of Observations by Age (Census 1911)



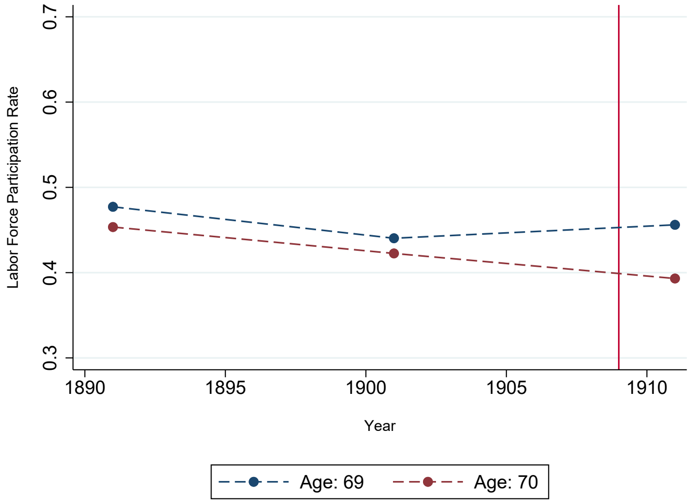
Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure 2: Labor Force Participation by Age



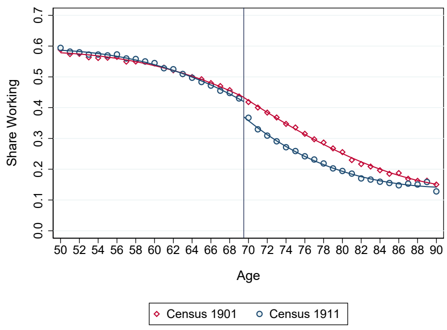
Source: Own calculations based on UK Census (waves 1901 and 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure 3: Labor Force Participation Over Time by Age-Based Eligibility

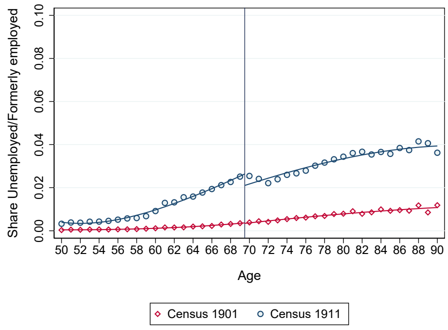


Source: Own calculations based on UK Census (waves 1891, 1901 and 1911) and IPUMS. *Note:* The vertical line indicates the introduction of old-age assistance by the OPA in 1909.

Figure 4: Work Status



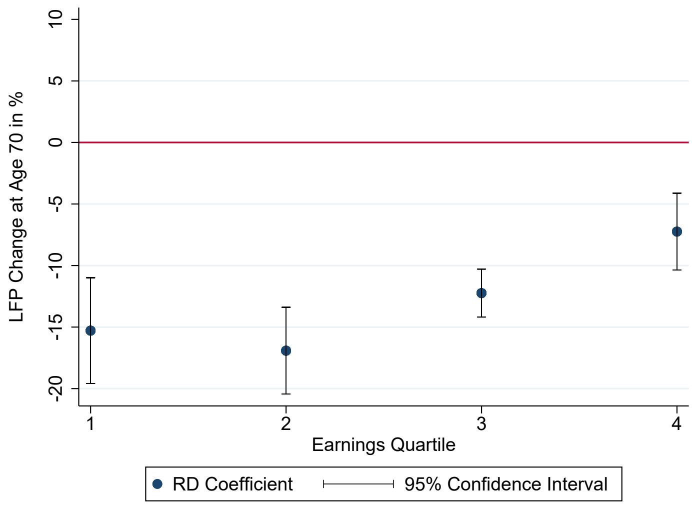
(a) Working



(b) Unemployed

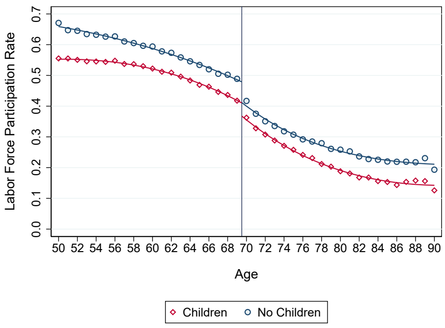
Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure 5: RD Estimates Across the Occupational Earnings Distribution

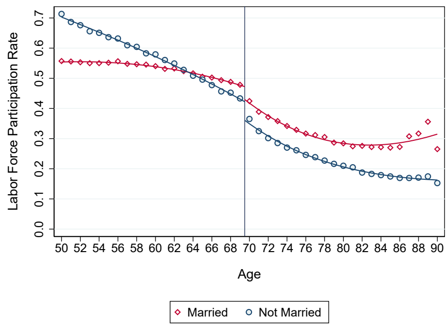


Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* RD estimates of LFP by earnings quartile in percent. For additional details see appendix C.

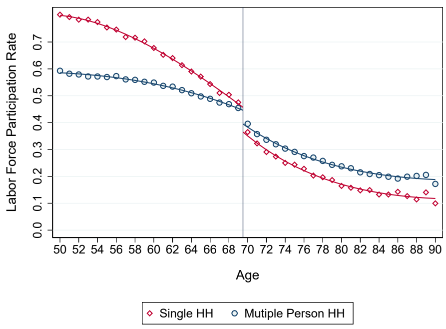
Figure 6: Labor Force Participation by Family Background (1911)



(a) Children



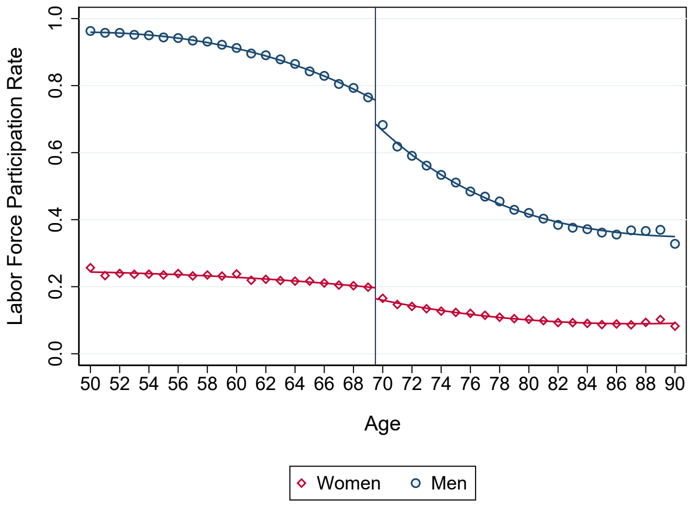
(b) Married



(c) Single Households

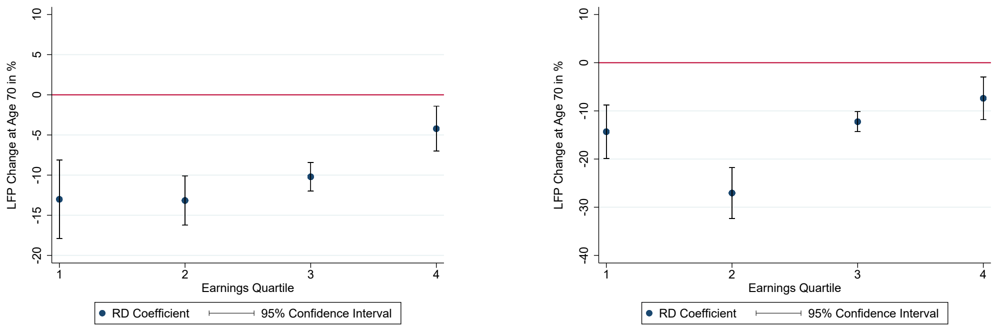
Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* Reported values are labor force participation rates over age. The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure 7: Labor Force Participation by Gender (1911)



Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* Reported values are labor force participation rates over age, separately for men and women. The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

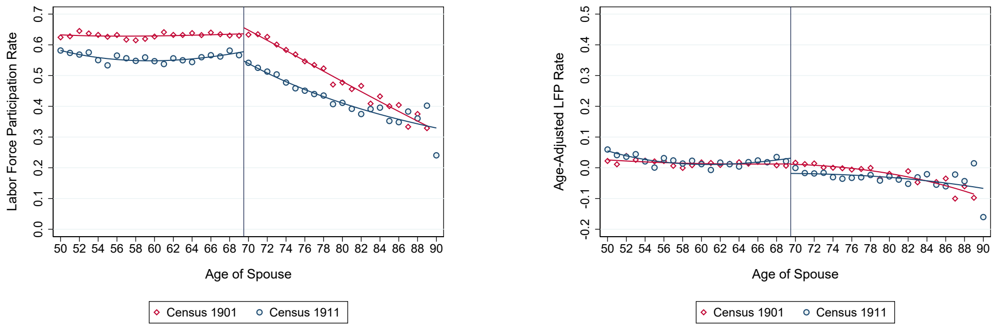
Figure 8: RD Estimates Across the Occupational Earnings Distribution by Gender



(a) Men (b) Women

Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* RD estimates of LFP by earnings quartile in percent, separately for men and women.For additional details see appendix C.

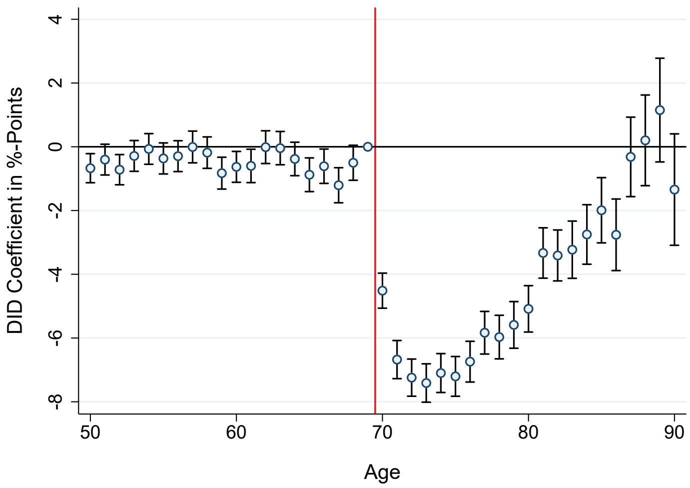
Figure 9: Labor Force Participation of Married Men by Age of Spouse



(a) LFP (b) Age-Adjusted LFP

Source: Own calculations based on UK Census (waves 1901 and 1911) and IPUMS. *Note:* LFP rates of married men (vertical axis) are conditional on having reached the eligibility age (71+) and are plotted by age of the respective spouse (horizontal axis). Panel (a) depicts LFP rates, panel (b) depicts LFP rates adjusted by own age. The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure 10: Extended DiD on Anticipatory Behavior



Source: Own calculations based on UK Census (waves 1901 and 1911) and IPUMS. *Note:* Reported DiD estimates (and 95% confidence bands) look at LFP changes from 1901 to 1911 for each age (50 to 90) relative to the reference age 69. The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

A Additional Tables (For Online Publication)

Table A.1: RD Estimates Native-Born vs. Foreign-Born

Sub-Sample	RD Coefficient	Observations
Native-Born	-0.061** (0.007)	1,247,484
Foreign-Born	-0.062** (0.005)	61,772

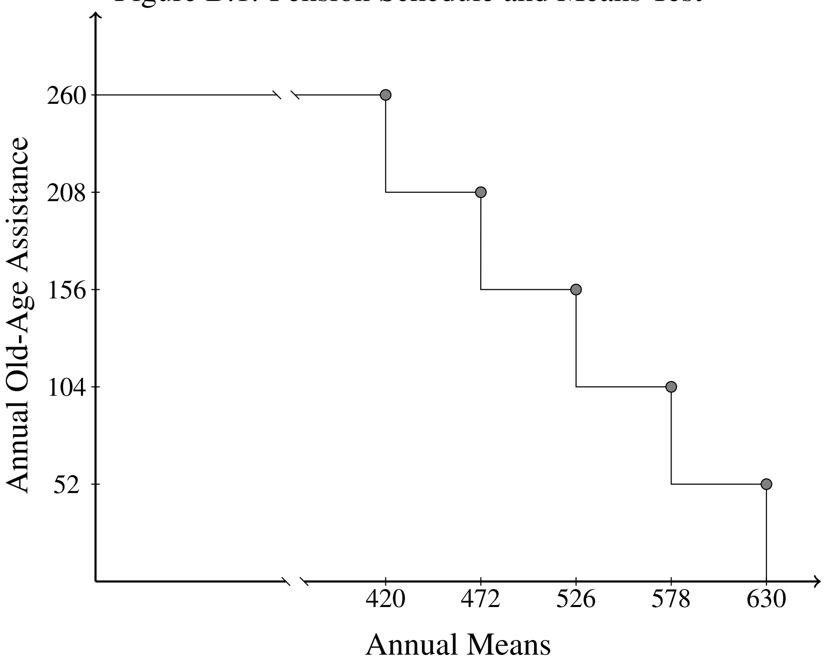
Source: UK Census wave 1911 and IPUMS. *Note:* RD estimates of the labor force participation rate (dependent variable) using an indicator for the age cutoff (= 0 if age < 70; = 1 if age >= 70). Estimates for sub-samples as indicated. Sum of native and foreign-born does not add up to the total number of observations from the complete sample, because the country of birth is unkonwn for 45,052 people. All regressions use a bandwidth of 5 age-years to the left (age 65-69) and to the right (age 70-74) of the cutoff, a linear polynomial in age, and uniform weighting on all observations. **, * denotes significance at the 1% and 5% level respectively. Standard errors (in parentheses) are clustered at age in years.

Table A.2: Assumptions of the Welfare Analysis

	Benchmark	Extreme Case
MVPF of Old-Age Assistance (WTP/Costs)	**0.80**	**0.37**
WTP for £1 of Old-Age Assistance	*0.91*	*0.76*
Counterfactual LFP Trend (71+)	Same as 65-69	Same as 65-69 + 10 PP
Net Cost of £1 of Old-Age Assistance	*1.13*	*2.03*
Administrative Cost to Mechanical Cost	3%	3%
Behavioral Cost to Mechanical Cost	10%	100%
MVPF of the Tax Increase	**1.02**	**1.95**
Pareto Parameter	1.5	1.5
Tax Rate (before)	3.8%	3.8%
Tax Rate (after)	8.3%	8.3%
Elasticity of Taxable Income	0.25	5
MVPF Ratio: $\frac{Tax}{Pension}$	**1.28**	**5.23**

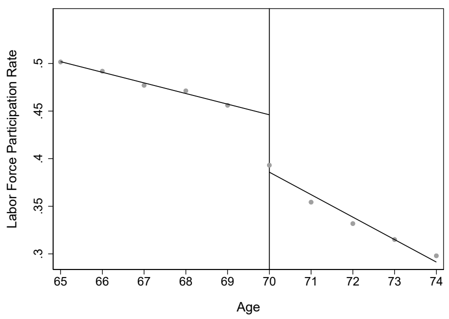
B Additional Figures (For Online Publication)

Figure B.1: Pension Schedule and Means Test

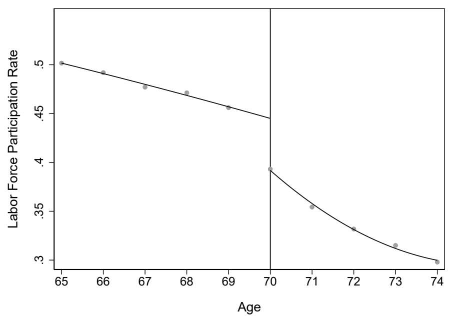


Source: Own graph.
Note: The figure depicts critical values of the means test. 20 Shillings = 1 Pound. The maximum pension of 260 Shillings per year (5 Shillings per week) was granted to individuals with annual means of no more than 420 Shillings (36% of the average annual earnings in 1911). Individuals with annual means of more than 630 Shillings (54% of the average annual earnings in 1911) did not qualify for the pension.

Figure B.2: Functional Form of Labor Force Participation and Age Around the Cutoff



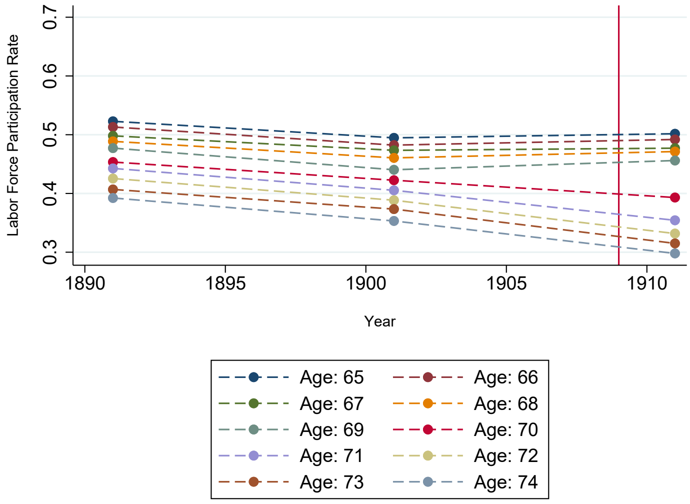
(a) Linear



(b) Quadratic

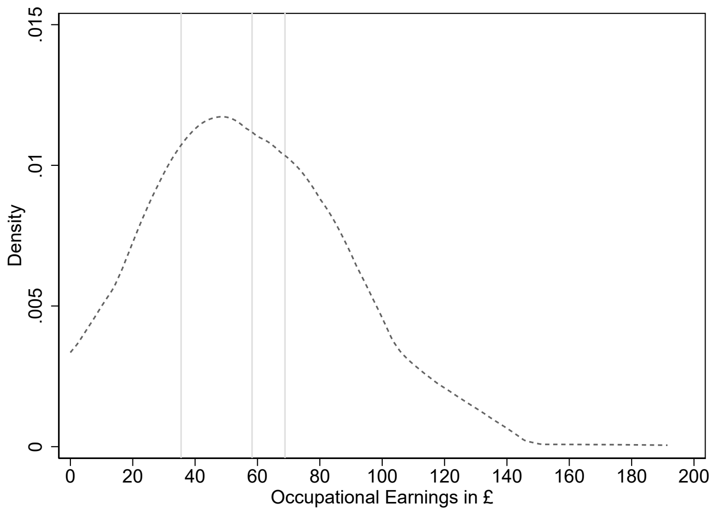
Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure B.3: Labor Force Participation Over Time by Age-Based Eligibility: More Ages



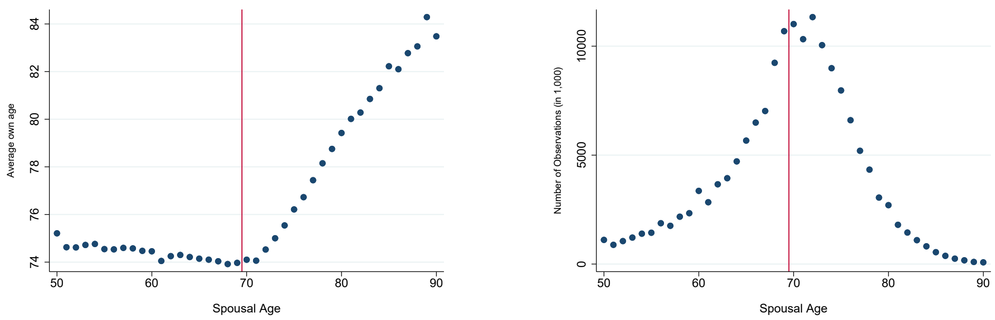
Source: Own calculations based on UK Census (waves 1891, 1901 and 1911) and IPUMS. *Note:* The vertical line indicates the introduction of old-age assistance by the OPA in 1909.

Figure B.4: The Distribution of Earnings in 1911



Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The graph reports the earnings distribution as approximated from 3-digit occupational scores (for details, see appendix C). Vertical lines indicate the earnings quartiles.

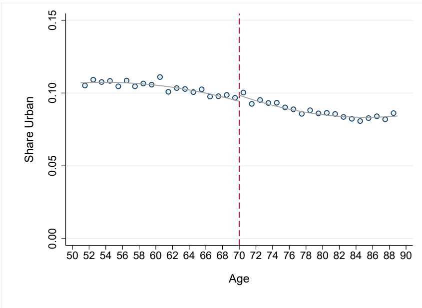
Figure B.5: Distribution of Age Differences Between Married Men and Their Spouse



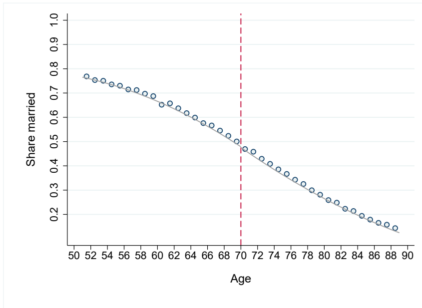
(a) Average Age by Spousal Age (b) Observations by Spousal Age

Source: Own calculations based on UK Census (waves 1901 and 1911) and IPUMS. *Note:* Reported values are conditional on men having reached the eligibility age (70) and are plotted by age of the respective spouse (horizontal axis). The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

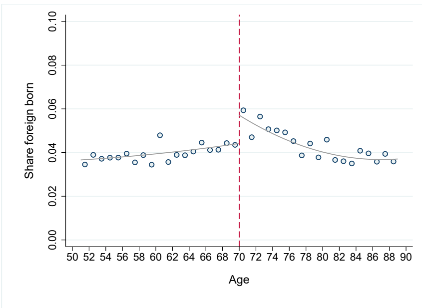
Figure B.6: Continuity of Observable Characteristics at the Age Cutoff



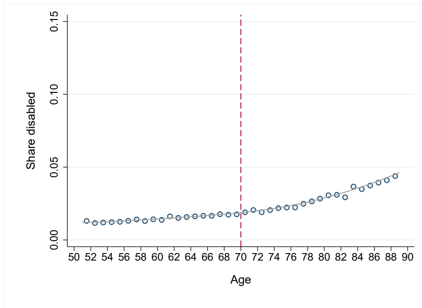
(a) Urban



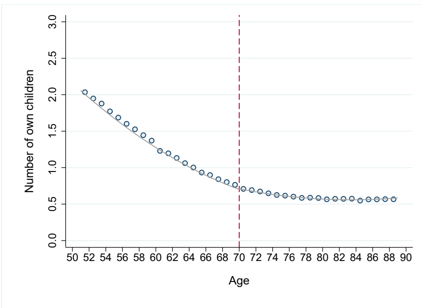
(b) Married



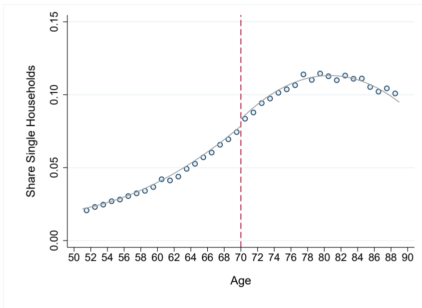
(c) Foreign born



(d) Disabled



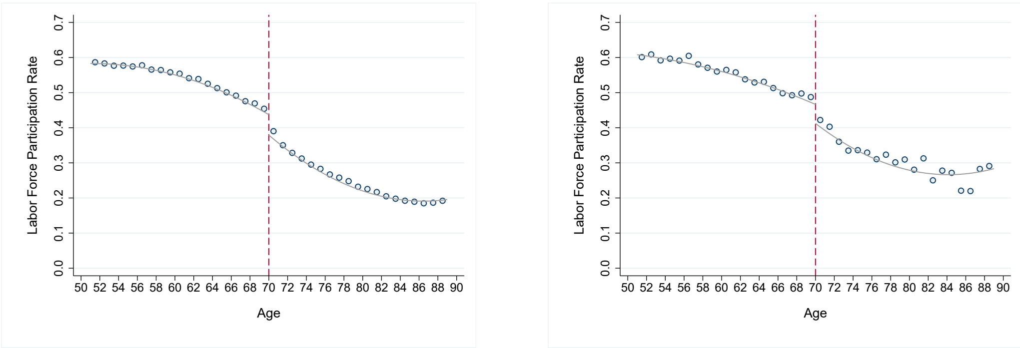
(e) Own Children in Household



(f) Share Single Households

Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

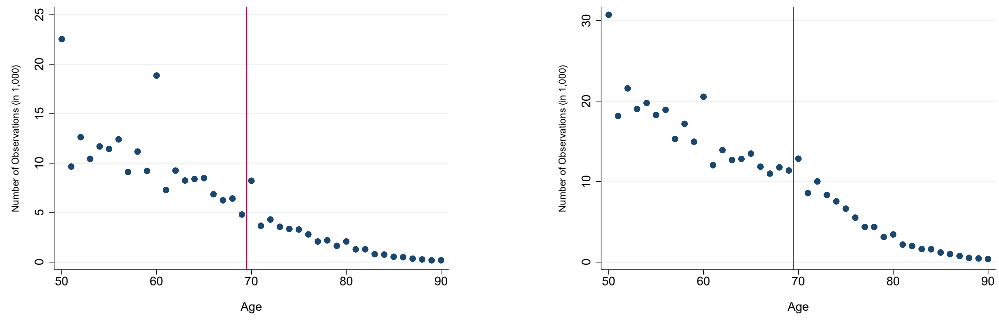
Figure B.7: Labor Force Participation and Native-Born and Foreign-Born



(a) Native-Born (b) Foreign-Born

Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

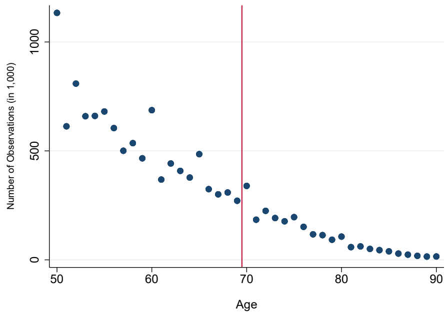
Figure B.8: Number of Foreign Born People by Age



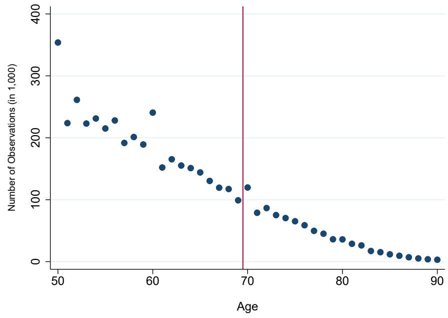
(a) 1901 (b) 1911

Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

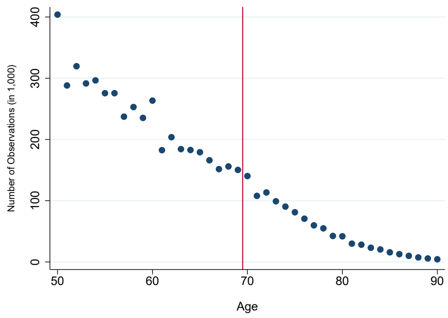
Figure B.9: Age-heaping in the US (1910) and the UK (1901) compared to the UK (1911)



(a) US 1910



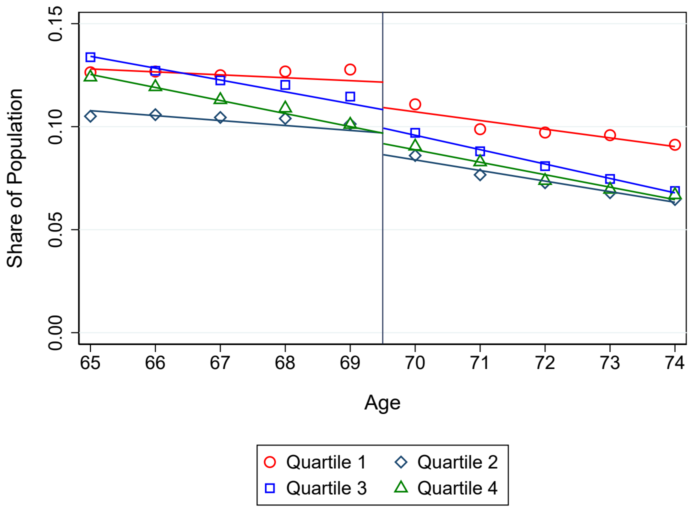
(b) UK 1901



(c) UK 1911

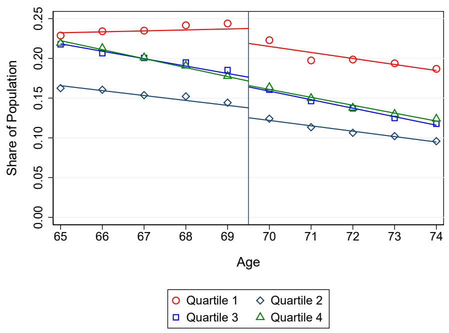
Source: Own calculations based on the full-count census data provided by IPUMS. *Note:* The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure B.10: Labor Force Participation by Earnings Group

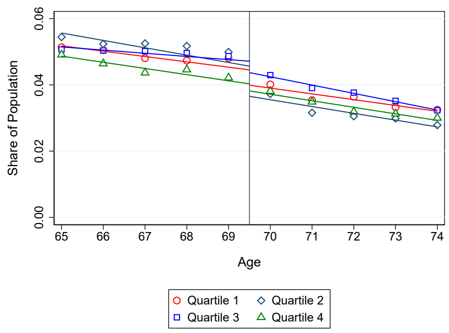


Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The graph reports the share of the population in every earnings quartile. Earnings are approximated from 3-digit occupational scores (details in appendix C). The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

Figure B.11: Labor Force Participation by Earnings Group



(a) Men



(b) Women

Source: Own calculations based on UK Census (wave 1911) and IPUMS. *Note:* The graph reports the share of the population in every earnings quartile. Earnings are approximated from 3-digit occupational scores (details in appendix C). The vertical line indicates the age-based eligibility threshold between age 69 and 70 that was introduced by the OPA in 1909.

C Approximating Occupational Labor Earnings (For Online Publication)

Earnings are not reported directly in the UK census waves used in this paper. We thus approximate earnings based on occupational status as has been done in the previous literature (e.g. Abramitzky et al., 2014). For this purpose, we match the occupational income score (occscore = median income by occupation) that is generated from the three-digit occupational level (occ1950) from IPUMS for the US census in 1950 to the UK census in 1911 (see details below). We use the US 1950 census because individual level census data that include earnings are not publicly available for the UK before 1991. Moreover, comprehensive earnings information for the US has not been collected before 1950.

Based on this procedure we can match an earnings score to 97% of individuals who state an occupation in the UK census of 1911. The implicit assumption behind this procedure is that the earnings distribution, conditional on the distribution of occupations, is similar in the UK in 1911 and the US in 1950. This assumption might be violated for country-specific (UK vs. US) or chronological (1911 vs. 1950) reasons. Given the lack of comparable income data for the UK at the time, it is difficult to assess the assumption empirically. However, the occupational income distribution that we derive from the US census of 1950 matches the main patterns of Feinstein (1990), who has calculated earnings for a range of occupations for the UK in 1911. In line with the results of Feinstein (1990), imputed earnings are lowest for farmers and domestic servants, followed by general laborers to miners or other skilled crafts. Merchants and academic occupations such as doctors or teachers have the highest occupational earnings. Although Feinstein (1990) does not report earnings for merchants and academic occupations, it seems plausible that they had relatively high incomes in the UK of 1911. All in all, we believe that the occupational earnings derived from the US census are suitable to approximate occupational earnings in the UK earnings of 1911.

Since estimating heterogeneous effects by earnings groups (4.2.1) only require a very broad approximation of the earnings distribution, we consider the results to be informative even if some occupations are misclassified. In an earlier version of this paper, we estimated LFP responses by each of the 10 occupational groups presented in table 2. The conclusions were similar to those presented in section 4.2.1, with a considerably negative correlation between skill level/earnings and the size of the LFP response to the OPA. For high-skilled (and most likely high paid) occupations such as legislators, managers or professionals we found little LFP declines while for agricultural workers or people in elementary occupations the LFP declines were large.

C.1 Further Details on the Matching of the Income Score from the US 1950 Census to the UK 1911 Census

The occupational classification in IPUMS differs between the US and UK census. In particular, the three-digit occupational system used in the US (1950 Census Bureau Occupational Classification system: occ1950) does not directly match with the five-digit occupational level used in the UK (Historical International Standard Classification of Occupations HISCO: occhisco). To solve this problem, we make use of the fact that the US census of 1880 from IPUMS includes both occupational coding systems (occ1950 and occhisco), which means that every individuals' occupation is coded both in the occ1950 and occhisco coding scheme. We thus match the US census from 1880 that also includes the occupational income score (occscore), derived from the incomes of the US census 1950, to the UK census of 1911. In case that multiple occ1950 codes are matched to one occhisco code, we assign the average occscore for each occhisco code (except for "occupational title unclassifiable", "ambiguous responses", "other non-occupational response" and "no occupation and not in universe/not applicable").

To circumvent several adjustments with respect to aggregate price changes (deflating from 1950 to 1911), differential trends in GDP growth between the US and the UK, and the overall conversion of US dollars to British pounds, we normalize the mean value of earnings (measured in 1950 US dollars) to match the UK mean earnings in 1911 that are documented in historical earnings data. This simple and transparent procedure preserves the ranks of the earnings distribution that is needed for the analysis (depicted in figure B.4).

C.2 Further Estimation Details

While the UK census documents occupations for the entire working population, occupational information is only available for a selected subset of retirees. Restricting the sample to those who reported their former occupation would likely result in an underestimated LFP response for low-earnings occupations, because more prestigious occupations are more likely to be reported. Consider the extreme case where none of the individuals in low-earnings occupations report their former occupation. In this case, the average LFP rate in low-earnings occupations would always be 1 irrespective of age, because people who retire from these occupations (hence switch LFP from 1 to 0) always drop from the universe of low-earnings occupations.

Instead of estimating the LFP drop conditional on the (former) occupation, our approach is to exploit the change in the share of people in one earnings group relative to the total population. Intuitively, instead of estimating the LFP drop at age 70 for the sub-sample of carpenters, we estimate the drop in the share of active carpenters in the total population at age 70. In this case, the drop at age 70 includes people that report to be former carpenters as well as people who do not report their former occupation. The advantage of this approach is that we include all retired individuals irrespective of their reporting status because we only need occupational data for older workers who are still active.

The estimates in figure 5 and figure 8 are based on the following procedure:

First, we generate four equally sized occupational earnings groups for all people with occupational earnings > 0. Therefore, we order people that are in the relevant age group for our main specification (65 to 74 years) according to their occupational earnings. We define that everybody who is not in the labor force (e.g. retired people) to have no occupational earnings, and thus not to be part of any occupational earnings group. We also define that the small share of people that report an active occupation but could not be matched to an occupational group in the US census (around 3%) not to be part of any earnings group. Dropping these unmatched occupations from the sample altogether does not affect the estimates. The resulting four earnings groups are not exactly of the same size, because some occupational groups (e.g. farmers and laborers) are too large to fit exactly into one quartile. Based on the imputation procedure, all members with the same occupations have the same earnings and therefore must belong to the same earnings group.

Second, we generate four separate endogenous variables, one for each earnings group. The resulting dummy variable indicates whether the respective individual belongs to that specific occupational earnings group ($Earnings - Group_{1-4} = 1$) or not ($Earnings - Group_{1-4} = 0$). In the final step, we run separate RD regressions for each of the four endogenous variables, investigating whether the share of people belonging to that occupational earnings group changes at age 70.

Figure B.10 and B.11 show how the share of each occupational earnings group changes at the cut-off. Note that the values on the y-axis are much lower than for the baseline RD graphs because instead of looking at the LFP rate, the y-axis represents the share of people within each occupational earnings group. Since earnings are only imputed for the subset of people in the labor force (LFP=1), the sum of the four earnings group shares approximate the aggregate LFP rate. One reason why the sum of the shares do not exactly add up to the aggregate LFP shares in figure 2 (46% at age 69, 40% at age 70) is the small share of individuals for whom we cannot impute earnings.

Since the interpretation of the estimates changes (because we measure the change in the share of people in an earnings group instead of the change in the average LFP rate), figures 5 and 8 report the relative activity decline for each occupational earnings group. Therefore, we divide the estimated decline in the share of people active in an occupational earnings group at the cutoff by the percentage of individuals active in an occupational earnings group at the age of 69. The estimates in figure 5 are thus compared with the relative LFP decline of 13% in the main specification (not the absolute decline of 6 percentage points). The weighted average of the estimated LFP drops in figure 5 thus exactly matches the 13% LFP decline derived from our main estimate.

D Welfare Analysis (For Online Publication)

Hendren and Sprung-Keyser (2020) show that the overall welfare implications of policies like the OPA, which are characterized by an equity-efficiency trade-off, can be assessed using the framework of the 'Marginal Value of Public Funds' (MVPF). On the spending side, the MVPF measures the amount of welfare benefits that can be delivered to policy beneficiaries per unit of government spending on the respective policy. On the revenue side, the MVPF quantifies how much monetary welfare benefits are lost by raising one unit of government revenue.

The welfare analysis of the OPA boils down to two questions. First, how much does it cost to transfer £1 of welfare benefits from a top income earner to a pensioner? Second, how much is a society willing to pay for such a transfer? As shown by Hendren and Sprung-Keyser (2020), we can answer the first question by comparing the MVPF of the old-age assistance program to the MVPF of the tax increase. Following this framework, we will argue that it was relatively cheap to transfer monetary welfare benefits to pensioners since behavioral responses were arguably modest relative to the scale of the program. The answer to the second question depends on the degree of redistribution entailed in the reform and the inequality aversion of the society. Assuming a preference for equality, the society was willing to pay a high price for the transfer given that the reform redistributed from the very top to the bottom of the earnings distribution. Thus, we conclude that the OPA increased overall welfare if the society was at least slightly averse to inequality.

To reach this conclusion, we have to make some assumptions. These assumptions are necessary to deal with the unknown counterfactual labor supply of elderly people in absence of the reform and the elasticity of taxable income to increased taxation. The overall conclusion, however, is robust to extreme changes in these assumptions. We start by first calculating the MVPF of the old-age assistance program and the tax increase. Then we compare the ratio of the MVPF to the ratio of social welfare weights derived from a utilitarian welfare function with diminishing marginal returns to consumption.

D.1 MVPF of Old-Age Assistance

In the following, we essentially borrow from the methodology outlined in Hendren (2016) and Hendren and Sprung-Keyser (2020). According to these papers, the MVPF is calculated by dividing the aggregate willingness to pay (WTP) for the policy by the net cost to the government.

Willingness to Pay: To calculate the WTP, we distinguish two classes of pension recipients. The first group consists of infra-marginal individuals, who do not alter their labor supply behavior in response to the reform. Infra-marginal individuals value their pension fully such that their willingness to pay equals the amount of pensions received.[41] The second group consists of marginal individuals who reduce their labor supply in response to the reform, either due to substitution effects (following from the means test) or income effects. The envelope theorem implies that marginal recipients are indifferent between a world with and without the OPA and therefore have a WTP of zero.[42] The overall WTP for one additional pound of benefits is therefore simply the ratio of infra-marginal recipients divided by the total number of pensioners.

We calculate the share of infra-marginal individuals using our reduced form RD estimate at the pension eligibility age. Furthermore, we have to make an assumption on the counterfactual labor supply at older ages. In the benchmark scenario, we assume that the LFP of older people (71+) would have evolved similarly to the LFP of people aged 65-69 who are not eligible for the pension. [43] In this case, the WTP for £1 of additional pension benefits is £0.91 (for details on the calculation see part D.4 below)[44] In addition to our benchmark scenario, we also construct

[41]Theoretically, the WTP for £1 of pension benefits might even exceed £1 if pensioners derive additional utility from the fact that the OPA provides insurance against longevity.

[42]Invoking the envelope theorem is only valid for marginal reforms. In the case of the OPA, which induced a discrete jump from no pension to "full" (minimum) pension, marginal individuals probably had a positive WTP. For the US, Fetter and Lockwood (2018) show that marginal individuals had indeed a WTP substantially above zero. We make the conservative assumption that marginal recipients in the UK had a WTP of zero.

[43]We consider this to be a rather conservative assumption, because LFP rates of people aged 71 and older have fallen more strongly than for 65-69 year old's in the pre-treatment period.

[44]Our WTP estimate of 0.91 is slightly below the 0.95 calculated in the US context by Fetter and Lockwood (2018). However, since we abstain from a structural model (in contrast to Fetter and Lockwood, 2018), we restrict the WTP for marginal individuals to zero to err on the conservative side. If we were to assume that the WTP among marginal individuals was roughly the same as in the US, the overall WTP in the UK would be 0.99 and thus be even higher than in the US. For background: Fetter and Lockwood (2018) report that 48% of the individuals are infra-marginal and that the remaining 52% are marginal individuals. A back-of-the-envelope suggests that the WTP among marginal individuals in the US was therefore roughly 0.9. Applying this number to the UK context yields a WTP of 0.99 ($1*0.91+0.9*0.09$).

an extreme case scenario (see Table A.2 for an overview of the assumptions) to challenge our overall conclusion. In that scenario, we assume that the LFP of older people (71+) relative to the age group 65-69 would have increased by 10 percentage points in absence of the reform which reduces the WTP to 0.76. Although this scenario is somewhat arbitrary and unrealistic, it is useful to show that our welfare statement remains robust even under extreme assumptions.[45]

Net Cost to the Government: The net costs of the policy to the government can be divided into mechanical, administrative and behavioral costs. Mechanical costs are simply the pension payments. For each additional £1 of pension payments, the government faces costs of £1. The administrative costs add up to about 3% of the mechanical costs of the program.[46] Behavioral costs might arise for two reasons: First, people retire due to the income effect before the eligibility age, leading to higher costs for other government programs. Second, people retire earlier (either before or after the eligibility age) which results in lower tax revenues. We argue below that both channels did play only a minor role (see part D.5 with details on behavioral costs). In the benchmark case, we conservatively assume that behavioral costs of old-age assistance were 10% of the mechanical costs of the reform (100% in the extreme case). In total, the net costs to the government add up to £1.13 for each £1 of pension payments (mechanical costs = 1, administrative costs = 0.03 and behavioral costs = 0.10).

MVPF: Our baseline calculation yields an MVPF of 0.80[47] (0.37 in the extreme case) which is very much in line with modern cash transfers analyzed in Hendren and Sprung-Keyser (2020). An MVPF of 0.80 implies that £1 of additional pension spending results in a £0.80 increase in monetary welfare benefits for pension recipients.

[45]Given the lack of data, we cannot make a statement on intensive margin adjustments (hours worked) in response to the reform. Given that most elderly people would have been out of the labor force even in absence of the pension, intensive margin adjustments for the remaining elderly in the labor force need to be massive to change the WTP meaningfully.

[46]Administrative costs of the OPA were low and only amounted to 3% (Old Age Pensions Committee, 1919), because the OPA mainly relied on existing institutions (post-offices etc., see Murray, 2009).

[47]$MVPF = \frac{WTP}{Costs} = \frac{0.91}{1.13} = 0.80$

D.2 MVPF of Raising Top Marginal Income Tax Rates

As shown by Hendren and Sprung-Keyser (2020), the MVPF of a tax reform can be calculated using the formula:

$$MVPF = \frac{1}{1 - \frac{t}{1-t} * a * e} \tag{2}$$

where a is the Pareto parameter of the income distribution, t is the tax rate and e is the elasticity of taxable income with respect to the 'keep' rate (1- t).

We know from Atkinson (2005) that the Pareto parameter at the top of the income distribution in the UK was around 1.5 at that time. Marginal tax rates at the top were low compared to today's standard (before the OPA: 3.8%, after the OPA at the very top: 8.3%).[48] To compute the MVPF we only need an assumption about the elasticity of taxable income. In the benchmark calculation, we use 0.25 which – according to Diamond and Saez (2011) – is "a mid-range estimate from the empirical literature" (p. 171). In the extreme case we use an elasticity of 5. Based on these inputs we obtain an MVPF of 1.02 (extreme case: 1.95), which means that the costs of transferring money from top income earners to the government budget were low. Note that the MVPF is small compared to the tax reforms considered in Hendren and Sprung-Keyser (2020) (ranging from 1.16 to infinity) for two reasons. First, inequality in the early 20th century UK was relatively high. Hence, raising the marginal tax rates at the top increased revenues substantially. Second, and more importantly, tax rates were comparatively low, thus making behavioral responses less likely.

D.3 Equity-Efficiency Trade-off

Dividing the MVPF of the tax increase by the MVPF of the old-age assistance program, yields 1.28 (extreme case: 5.23). This ratio tells us that the cost of transferring £1 of welfare benefits

[48] We follow Hendren and Sprung-Keyser (2020) and calculate the fiscal externality $-\frac{t}{1-t} * a * e$ by taking the average for the new and old tax rate.

to a pensioner via an increase of the top marginal income tax rate was £1.28, which is cheaper than transferring 1$ from the top to the bottom of the US income distribution in the current US tax system (which would cost 1.5$ - 2$ according to Hendren, 2020). The lower costs in the case of the OPA are especially driven by the small behavioral costs of increasing marginal top income tax rates.

To answer the question whether the society was willing to pay this price for the transfer, we need knowledge on the societies' value of £1 of additional welfare benefits in the hands of a pensioner, relative to a top income earner. Such a statement requires an assumption about the social welfare function. For illustrative purposes, we adopt a standard utilitarian welfare function (overall welfare equals the unweighted sum of individual welfare) based on a homogeneous utility function with diminishing returns to consumption (a log-utility function). In this case, the ratio of the social welfare weights for two population subgroups should equal the ratio of their marginal utility of consumption (Finkelstein, 2019). Dividing the marginal utility of consumption at the earnings region where the pension means-test phases in (£30) by the marginal utility of consumption at the top of the earnings distribution, where top marginal income taxes are levied (£5000), yields a ratio of 164.[49] This ratio implies that, based on a log-utility welfare function, the society would have been willing to spend up to £164 for a £1 welfare transfer compared to the actual cost of £1.28 (or in the extreme case: £5.23). Although the specific functional form of the social welfare function is debatable, our results show that the OPA was welfare enhancing even if the society had only a very small preference for equality.

D.4 Details on the WTP for the Pension

We calculate the overall WTP for one additional pound of pension benefits by dividing the ratio of infra-marginal recipients by the total number of pensioners. To compute the share of infra-marginal individuals, we need to know how many pension recipients retired in response to the OPA. At the eligibility age cutoff, our RD estimate provides a good indication about the retirement response. According to our baseline estimate, LFP declined by 6.0 percentage points at the cutoff, implying that approximately 8,459 persons at age 70 (6% of 140,288) retired due

[49] Using the lowest income region where marginal tax rates were increased (£2000) yields a ratio of 66.

to the OPA. To calculate how many people retired at older ages (71+), we need an assumption on the counterfactual LFP of older people in absence of the reform. In the benchmark scenario, we assume that the LFP of older people would have evolved similarly to the LFP of the age group 65-69 who are not eligible for the pension. Under this assumption, the LFP rate of the age group 71+ in 1911 would have been 33.0% in contrast to the actual observed LFP rate of 27.9%.[50] The counterfactual LFP rate implies that 47,909 people aged 71 and older ((33.0%-27.9%) $*$ 928,198[51]) retired as a response to the OPA. The total number of retirees at the age cutoff (8,459) and at later ages (47,909) yields 56,368 marginal individuals. Given the total number of pension recipients in 1911 was 613,873, the WTP for £1 more of pension benefits would have been £0.91 ($\frac{613{,}873-56{,}368}{613{,}873}$).

In the extreme case scenario, we assume that the LFP of older people (71+) relative to the age group 65-69 would have increased by 10 percentage points in absence of the reform. This would yield an counterfactual LFP rate of 43% ($32.2+0.8+10=43.0$) among people aged 71+ so that the WTP for the OPA declines to 0.76.

D.5 Details on the Behavioral Costs of the Pension

Behavioral costs of the pension for the state budget can arise for two reasons: First, people retire due to the income effect before the eligibility age, leading to higher costs for other government programs. Second, people retire earlier (either before or after the eligibility age) which results in lower tax revenues. The first channel was negligible in practice because other social programs – except for the pension – were either non-existent at the time (there was no unemployment or health insurance in place) or attached to tough conditions (e.g. the obligation to work in a poor house) and social stigma (including the loss of voting rights). Hence, even if people retired prior to the eligibility age, which does not seem to be the case (see section 4.3.4), they would probably not have relied on other government programs. Tax revenues were also hardly affected by behavioral changes of older workers because their tax burden was low in the first

[50]If the age-specific LFP had remained on its 1901-level, then 32.2% in the age group 71+ would have been in the labor force in 1911. Among those aged 65-69, the LFP rate increased by 0.8 percentage points between 1901 and 1911. Thus the (unknown) counterfactual in the group 71+ adds up to $32.2+0.8=33.0$.

[51]928,198 marks all individuals alive at ages above 70.

place. Individuals earning below £160 did not pay any income tax (House of Commons, 1911), which was more than 2.7 times the mean earnings in 1911 (£58.6, Feinstein, 1990). Thus, only few individuals were subject to income taxes in the UK at that time.[52] The tax burden among the elderly was even lower given that they earned much less than the average citizen (Spender, 1892).

[52] There were only about 1 million income tax payers in the UK in 1909/1910 compared to 36,353,455 inhabitants in England and Wales alone in 1911.