Macroeconomic Dynamics, **24**, 2020, 1635–1673. Printed in the United States of America.
doi:10.1017/S1365100518000962

ECONOMIC-DEMOGRAPHIC DEPENDENCY RATIO IN A LIFE-CYCLE MODEL

SAU-HIM P. LAU
University of Hong Kong

ALBERT K. TSUI
National University of Singapore

The conventional dependency ratio based on cohort-invariant cutoff points could overstate the true burden of population aging. Using optimal cohort-varying years of schooling and retirement age in a life-cycle model, we propose a modified definition of dependency ratio. We compare the proposed economic-demographic dependency ratio (EDDR) with the conventional definition and find that the conventional dependency ratio of the USA is projected to increase by 0.105 from 2010 to 2060, which is an over-projection of 86% when compared with the projected increase of 0.015 in the EDDR over the same period. Sensitivity analysis suggests that our finding is quite robust to reasonable changes in parameter values (except for one parameter), and the magnitude of over-projection ranges mainly from 0.079 to 0.102 (i.e., 75% to 97%). We follow the well-established Lee–Carter model to forecast stochastic mortality and employ the method of expanding duration to decompose the sources of over-projection.

Keywords: Rapid Population Aging, Behavioral Response, Economic-Demographic Dependency Ratio, Conventional Dependency Ratio

1. INTRODUCTION

One of the statistics most commonly used to measure population aging is the dependency ratio.[1] Total dependency ratio as well as its components (child and old-age dependency ratios) of USA from 2010 to 2060, are shown in Figure 1.[2] It is observed that total dependency ratio increases from 0.67 to 0.87 in 50 years,[3] meaning that 100 working people in 2060 will have to support 20 more young or old people, as compared to those in 2010.

Because of substantial fertility and mortality reductions anticipated in the coming decades for most advanced countries, it is not surprising that the dependency

We thank the University of Hong Kong (Project No. 104003954) for financial support. We are grateful to Kelvin Yuen for valuable research assistance, and to an anonymous referee and the participants of the Asian Meeting of the Econometric Society (Hong Kong), Annual Conference of the European Society for Population Economics (Glasgow) and 25th Annual Colloquium of Superannuation Researchers (Sydney) for useful comments. Address correspondence to: Sau-Him P. Lau, Faculty of Business and Economics, University of Hong Kong, Hong Kong. e-mail: laushp@hku.hk. Phone: (852) 2857-8509. Fax: (852) 2548-1152.

© 2019 Cambridge University Press 1469-8056/19

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

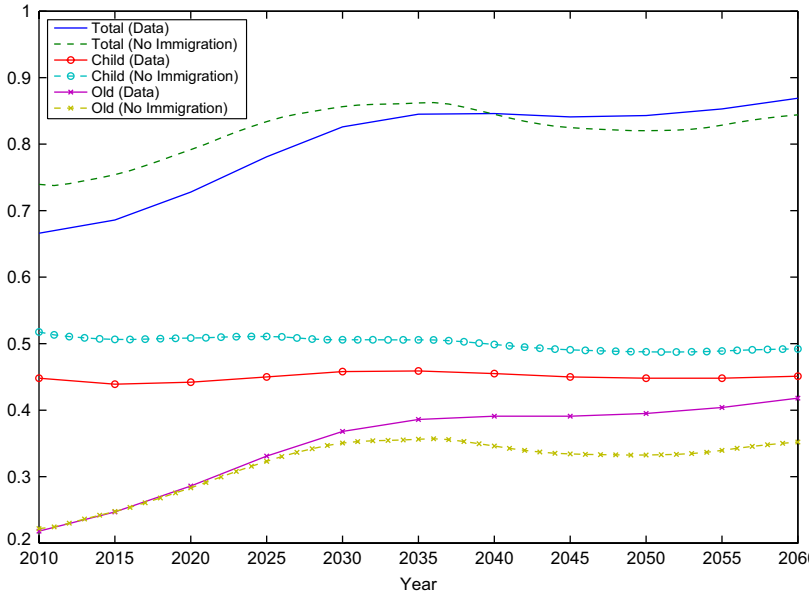


FIGURE 1. Conventional dependency ratios.

ratios, when defined in the conventional way, exhibit huge changes. However, do changes in these statistics represent accurately the changing trend of dependency of "resource consumers" upon "resource producers"? An implicit assumption of the conventional dependency ratio is that a person is considered as young (and needing resource support) before age 20 and as elderly (and needing resource support too) after 65. We think that the *cohort-invariant cutoff points* in the conventional definition need to be examined carefully. While it may be reasonable to assume that people who aged between 20 and 65 in a particular year (say, 1950) are resource producers and those outside this range are resource dependents, it is less clear that the same cutoff points also lead to an accurate description of *economic resource dependency* at a different point of time (say, 50 years later), when people generally live longer and healthier.

Researchers from various academic disciplines have pointed out that the conventional definition of dependency ratio may need refinement in the presence of substantial demographic changes. Some researchers, such as Sanderson and Scherbov (2005) and Lutz et al. (2008), argue that several demographic measures, including the median age and the dependency ratios, are based on years since birth (i.e., chronological age) only. To improve on these purely backward-looking measures, they propose *forward-looking demographic* concepts (such as the prospective old-age dependency ratio (POADR)) to adjust for life expectancy increase.[4] They also show that from 2005–10 to 2045–50, the POADR increases less rapidly (by 54%, from 0.13 to 0.20) than the conventional old-age dependency ratio (by 81%, from 0.21 to 0.38) in USA.

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

While the forward-looking POADR may represent a useful alternative to the conventional measure, we suggest in this paper that the purely demographic definition could be improved further by incorporating appropriate economic ingredients. In the current context of facing rapid population aging, possible responses in important economic variables such as age of retirement [Bloom et al. (2007) and d'Albis et al. (2012)] and years of schooling [Hazan (2009) and Cervellati and Sunde (2013)] are not incorporated in the conventional definition of dependency ratio, as well as the modified definition suggested in Sanderson and Scherbov (2010). This paper proposes a modified definition of dependency ratio, to be called the "*economic-demographic* dependency ratio" (EDDR). Similar to POADR, it has the forward-looking dimension, but instead of using a purely demographic definition, the forward-looking element in our proposed definition is based on behavioral response guided by a well-articulated economic model.

Our proposed definition of EDDR is based on a life-cycle model with *schooling and retirement choices*. We first obtain the first-order conditions characterizing the behavior of individuals from different cohorts. We then perform computational analysis and find that optimal years of schooling and age of retirement increase over time in this model. Using the well-known method of mortality forecasting by Lee and Carter (1992) and ignoring immigration for easy comparison, we find that the EDDR increases by 0.015 from 2010 to 2060 in USA. We compare our results with the conventional dependency ratio, also without immigration,[5] and find that the conventional dependency ratio increases by 0.105 during this period, implying that it over-projects by 0.09 (or 86%) when compared with the EDDR. We conduct sensitivity analysis and find that out of the increase of 0.105 in conventional dependency ratio from 2010 to 2060, 0.079 to 0.102 are likely due to over-projection because the conventional dependency ratio does not incorporate behavioral response to demographic changes. In summary, our analysis suggests that if individuals respond to population aging by adjusting schooling and retirement behavior according to the life-cycle model, then the conventional dependency ratio (which does not reflect these behavioral changes) may over-project the impact of population aging in the next few decades.

Besides the EDDR suggested in this paper, alternative definitions of dependency ratio have been proposed in the literature. A detailed discussion of the various approaches can be found in Sanderson and Scherbov (2015) and Loichinger et al. (2017). We highlight two related approaches using economic concepts to modify the conventional dependency ratio. First, the concept "economic dependency ratio" has been used in official publications such as the Monthly Labor Review. It incorporates both extensive and intensive margins of labor force participation [Sanderson and Scherbov (2015, pp. 692–694) and Loichinger et al. (2017, Section 2.2)]. Second, the National Transfer Accounts (NTA) economic support ratio has also been widely used [Sanderson and Scherbov (2015, pp. 694–698) and Loichinger et al. (2017, Section 2.3)]. This approach focuses on transfer of resources across generations and to and from the government and is based on the results developed in the NTA project [Lee and Mason (2011) and

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

Lee and Mason (2013)]. While these two approaches and the EDDR share similarities in that they are mainly based on economic concepts, there are also differences among them. In particular, the EDDR is different from these two approaches in at least two aspects: our proposed definition is more explicitly based on a simple but well-articulated economic model and our choices of the underlying ingredients of the model are as transparent and easily replicable as possible. We will elaborate on these points in the Conclusion, after presenting our method and results.

This paper is organized as follows. In Section 2, we introduce the economic model and describe how the EDDR is constructed. We conduct quantitative analysis on the EDDR from 2010 to 2060 in Section 3 and compare it with the conventional definition in Section 4. We perform sensitivity analysis in Section 5. Concluding remarks are offered in Section 6.

2. A NEW DEFINITION OF DEPENDENCY RATIO

The conventional dependency ratio, which is based on cohort-invariant cutoff points and does not include changes in behavior, may not capture the impact of population aging accurately. We aim to provide an improved definition by incorporating behavioral response. To contrast the proposed behaviorally based definition with the conventional dependency ratio as sharply as possible, we make two simplifications in the following analysis. First, we focus on the extensive margin, but not the intensive margin, of labor supply. We think that the simple and transparent features of the conventional definition are advantages that should be preserved as much as possible in the new definition. Thus, we aim to develop an alternative definition with the possibility of cohort-varying cutoff points, while keeping other aspects of the original definition, especially that the underlying age-specific labor force participation rate is either 0 or 1, unchanged.[6] To achieve this feature, we consider an economic model with individual's birth year as the only source of heterogeneity and focus on two key life-cycle decisions: schooling versus entering the job market and continuing working versus retiring. Second, we consider an environment without immigration and only focus on fertility and mortality factors. While it is possible to extend our definition to incorporate immigration data, we decide not to consider this factor because the key difference between the two definitions of dependency ratio does not depend on whether immigration is present or not.

2.1. Schooling and Retirement Choices

We consider a continuous-time life-cycle model with uncertain lifetime. We keep as many standard features as possible in the model. An individual chooses her consumption path (up to Ω, the maximum biological age), years of schooling and age of retirement. The lifespan uncertainty of cohort-b individuals is represented by a general survival function $\widetilde{l}_b(x)$, where $\widetilde{l}_b(x)$ is the probability that a cohort-b

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

individual survives up to at least age x. The survival function, which is allowed to be cohort-specific, reflects the trend of mortality decline.

It is assumed in this model that individuals make economic decisions starting from age M. For the theoretical analysis, it is helpful to focus on "adult age" (i.e., age measuring from age M). Conditional on reaching age M, the probability that a cohort-b individual survives up to at least age x (i.e., actual age $M+x$) is denoted by $l_b(x)$, where $x \in [0, \Omega - M]$.[7] This conditional probability, defined on adult age, is related to the original survival function $\widetilde{l}_b(x)$, defined on actual age, according to

$$l_b(x) = \frac{\widetilde{l}_b(x+M)}{\widetilde{l}_b(M)}. \tag{1}$$

We assume that schooling and labor supply choices are indivisible and that the progression from schooling to working and then to retirement are irreversible. These assumptions, which have been used by many researchers such as Kalemli-Ozcan et al. (2000), Hazan (2009), and Sánchez-Romero et al. (2016), will generate the resulting EDDR in a form similar to that of the conventional dependency ratio, except for the cohort-invariant cutoff points.

In this environment, an individual chooses a consumption path, years of schooling (S), and age of retirement (R) to maximize expected lifetime utility

$$\int_0^R e^{-\rho x} l_b(x) \left[\frac{c(x)^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] dx + \int_R^{\Omega - M} e^{-\rho x} l_b(x) \left[\frac{[(1+\theta)\, c(x)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] dx, \tag{2}$$

subject to

$$a'(x) = \begin{cases} [r + \mu_b(x)]\, a(x) + \xi_b h(S) - c(x) & \text{if } S < x \le R \\ [r + \mu_b(x)]\, a(x) - c(x) & \text{if } x \le S \text{ or } x > R \end{cases}, \tag{3}$$

and boundary conditions $a(0) = 0$, $a(\Omega - M) \ge 0$, where ρ is the discount rate, σ is the coefficient of intertemporal elasticity of substitution, θ is a parameter capturing the utility from leisure during retirement, r is the constant real interest rate, ξ_b is the index of productivity level of a cohort-b individual, $h(S)$ is the human capital level of the individual, $c(x)$ is consumption at age x, $a(x)$ is financial wealth at age x, $a'(x)$ is the derivative of $a(x)$ with respect to x, and

$$l_b(x) = e^{-\int_0^x \mu_b(q) dq}, \tag{4}$$

where $l_b(0) = 1$, $l_b(\Omega - M) = 0$, and $\mu_b(q) \ge 0$ (with $\lim_{q \to \Omega - M} \mu(q) = \infty$) is the instantaneous mortality rate at age q. We assume that $\rho \ge 0$, $r \ge 0$, $0 < \sigma \le 1$, and $\theta > 0$.

Unlike the disutility of labor specification used in many life-cycle models with endogenous age of retirement [such as Hazan (2009), d'Albis et al. (2012), and Bloom et al. (2014)], we follow Manuelli et al. (2012) to assume that the form of instantaneous utility function (2) in the retirement phase is different from that

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

during the schooling and working phases, so as to capture the drop in consumption level at retirement.

According to the budget constraint (3), when an individual works (after studying for S years), her wage rate is given by $\xi_b h(S)$. One may think of this specification as consisting of three components: depending on (a) an index ξ_b capturing the changing level of productivity of different cohorts, with a person from more recent cohort benefiting from a higher value of ξ_b, (b) one's level of human capital $h(S)$, and (c) the compensation to raw labor, which is normalized to be 1. Following the idea in the literature [such as Ben-Porath (1967), Hazan (2009), and Cervellati and Sunde (2013)], we assume that schooling has a productivity-enhancing role and the return to schooling, $\frac{h'(S)}{h(S)}$, is positive but decreasing in S.

The above specification reflects that the agent has no bequest motive, and a perfect annuity market exists to fully insure against mortality risk, similar to Yaari (1965). Therefore, at each age x, the agent can lend or borrow in a perfect financial market with effective (instantaneous) rate of return $r + \mu_b(x)$.

The individual's various choices are obtained as follows. (See Appendix B for detailed analysis.) First, conditional on a particular length of the schooling period and age of retirement, we obtain the optimal consumption choice of a cohort-b individual at age x, defined as $c_b(x, S, R)$. It can be shown that the (conditional) optimal consumption path is characterized by

$$c_b(x,S,R)=\begin{cases} e^{\sigma(r-\rho)x}c_b(0,S,R) & \text{if } x\le R \\ (1+\theta)^{\sigma-1}\,e^{\sigma(r-\rho)x}c_b(0,S,R) & \text{if } x>R \end{cases}, \tag{5}$$

where the initial consumption level, $c_b(0, S, R)$, is given by

$$c_b(0,S,R)=\frac{\xi_b h(S)\int_S^R e^{-rx}l_b(x)\,dx}{\int_0^R e^{-[(1-\sigma)r+\sigma\rho]x}l_b(x)\,dx+(1+\theta)^{\sigma-1}\int_R^{\Omega-M} e^{-[(1-\sigma)r+\sigma\rho]x}l_b(x)\,dx}. \tag{6}$$

Second, after substituting the optimal consumption path (5) into the objective function, we obtain the first-order conditions for interior optimal years of schooling (S_b^*) and age of retirement (R_b^*) as[8]

$$\xi_b h'(S_b^*)\int_{S_b^*}^{R_b^*} e^{-rx}l_b(x)\,dx=e^{-rS_b^*}l_b\left(S_b^*\right)\xi_b h(S_b^*), \tag{7}$$

and

$$c_b(0,S_b^*,R_b^*)^{-\frac{1}{\sigma}}e^{-rR_b^*}l_b\left(R_b^*\right)\left\{\xi_b h\left(S_b^*\right)-\left[1-(1+\theta)^{(\sigma-1)}\right]c_b(R_b^*,S_b^*,R_b^*)\right\}$$

$$=e^{-\rho R_b^*}l_b\left(R_b^*\right)\left[(1+\theta)^{(\sigma-1)}-1\right]\frac{c_b(R_b^*,S_b^*,R_b^*)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}, \tag{8}$$

where

$$c_b(R_b^*,S_b^*,R_b^*)=\lim_{x\to(R_b^*)^-}c_b(x,S_b^*,R_b^*)=e^{\sigma(r-\rho)R_b^*}c_b(0,S_b^*,R_b^*).$$

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

The first-order condition for optimal years of schooling is interpreted as follows. The left-hand side of (7) is the marginal benefit of continuing to study, which is measured by the expected present discounted value (PDV) of the increases in labor income throughout the working years from age S_b^* to age R_b^*, due to more years of schooling (and thus, higher level of human capital). The right-hand side of (7) is the marginal cost (the expected PDV of foregone labor income at age S_b^*) of postponing entry into the labor market.

The first-order condition for optimal age of retirement is more complicated for the life-cycle model with utility jump at retirement. The right-hand side of (8) is the marginal cost of delaying retirement. As seen from (A7b) in Appendix B, this cost is measured by the "utility jump at retirement" term, since the utility level at age R_b^* decreases when age of retirement increases (by an infinitesimal amount). On the other hand, the left-hand side of (8) is the marginal benefit of delaying retirement. By delaying retirement, one's labor income increases, but the consumption level (at R_b^*) is also higher (than the corresponding level if one retires). Thus, the marginal benefit (in monetary unit) is given by the difference of these two terms. After converting these two terms in monetary unit into utility level through multiplying by $c_b(0, S_b^*, R_b^*)^{-\frac{1}{\sigma}}$, the marginal utility of consumption at age 0, we obtain the marginal benefit of delaying retirement on the left-hand side of (8).

Under the assumption of a constant-intertemporal-elasticity-of-substitution (CIES) utility function (for both pre- and post-retirement phases) in (2), the first-order condition for retirement can be further simplified. Specifically, after substituting (5) and (6) into (8), we obtain

$$\frac{1-\sigma}{1-(1+\theta)^{\sigma-1}} = \frac{e^{\sigma(r-\rho)R_b^*}\int_{S_b^*}^{R_b^*} e^{-rx} l_b(x)\,dx}{\int_0^{R_b^*} e^{-[(1-\sigma)r+\sigma\rho]x} l_b(x)\,dx + (1+\theta)^{\sigma-1}\int_{R_b^*}^{\Omega-M} e^{-[(1-\sigma)r+\sigma\rho]x} l_b(x)\,dx}, \quad \textbf{(8a)}$$

and it is clear that ξ_b does not appear in (8a). We also observe that ξ_b appears in both the left- and right-hand sides of (7) and can be cancelled out. Therefore, it can be concluded that the productivity level ξ_b affects neither the first-order condition for years of schooling (S_b^*) nor that for age of retirement (R_b^*) in this economy.[9]

2.2. Dependency Ratios Based on the Life-cycle Model

The proposed EDDRs are based on S_b^* and R_b^* of the life-cycle model discussed in the previous subsection. Since the conventional dependency ratios are expressed in actual age, for easy comparison, we now convert relevant variables back to actual age (from adult age in the theoretical model).

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

For cohort-b individuals, define the working phase as between age $S_b^* + M$ to age $R_b^* + M$, the pre-working phase as between age 0 to age $S_b^* + M$, and the post-retirement phase as between age $R_b^* + M$ to age Ω. Based on this classification, define the following indicator functions for an individual born at time b and currently aged $t - b$:

$$I_{WA}(t,b) = \begin{cases} 1 & \text{if } S_b^* + M < t - b \leq R_b^* + M \\ 0 & \text{otherwise} \end{cases}, \tag{9}$$

$$I_{PrW}(t,b) = \begin{cases} 1 & \text{if } t - b \leq S_b^* + M \\ 0 & \text{otherwise} \end{cases}, \tag{10}$$

and

$$I_{PoR}(t,b) = \begin{cases} 1 & \text{if } t - b > R_b^* + M \\ 0 & \text{otherwise} \end{cases}. \tag{11}$$

The number of working-age persons at time t is given by

$$N_{WA}(t) = \int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{WA}(t,b)db, \tag{12}$$

where $B(b)$ is the number of births at time b. Similarly, the number of pre-working persons at time t is given by

$$N_{PrW}(t) = \int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{PrW}(t,b)db, \tag{13}$$

and the number of post-retirement persons at time t is given by

$$N_{PoR}(t) = \int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{PoR}(t,b)db. \tag{14}$$

Based on (12)–(14), we define the pre-working dependency ratio as

$$EDDR_{PrW}(t) = \frac{N_{PrW}(t)}{N_{WA}(t)} = \frac{\int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{PrW}(t,b)db}{\int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{WA}(t,b)db}, \tag{15}$$

and the post-retirement dependency ratio as

$$EDDR_{PoR}(t) = \frac{N_{PoR}(t)}{N_{WA}(t)} = \frac{\int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{PoR}(t,b)db}{\int_{t-\Omega}^{t} B(b)\, \widetilde{l}_b(t-b)\, I_{WA}(t,b)db}. \tag{16}$$

Finally, the total dependency ratio is simply the sum of these two ratios:

$$EDDR(t) = EDDR_{PrW}(t) + EDDR_{PoR}(t). \tag{17}$$

2.3. Utility Jump at Retirement Versus Disutility of Labor Specifications

In the life-cycle model described above, we use the utility jump at retirement specification [as in Manuelli et al. (2012)], contrasting with the disutility of labor

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

specification used in many papers [d'Albis et al. (2012), Bloom et al. (2014), and Sánchez-Romero et al. (2016)]. In this section, we compare these two specifications. In particular, we point out a different implication between them, which explains why we choose the utility jump at retirement specification in this paper.

These two specifications represent different ways to explain the retirement behavior. In considering whether to retire or not at a particular age, a benefit of extending the working duration is the labor income earned. If there is no cost in continuing to work, then an individual will always extend the working duration. The presence of the cost is to rationalize why people do not usually work until death. The disutility of labor approach captures the cost as a disutility term if one works, and usually this function is assumed to be non-decreasing in age. The utility jump at retirement specification captures the cost in a different way by assuming that the individual's utility derived from consumption (at an unchanged level) is lower when one works than when one is retired.

To see the different implications of these two specifications in the life-cycle model, we examine the first-order condition of retirement. For existing life-cycle models adopting the disutility of labor specification, the marginal cost of delaying retirement is given by the disutility of labor, as in Bloom et al. (2014) or Sánchez-Romero et al. (2016).[10] For the utility jump at retirement specification, the first-order condition is given by (8) above. The marginal cost of delaying retirement is now captured by the "utility jump at retirement." Moreover, this specification leads to an extra "consumption drop at retirement" term in the marginal benefit.[11]

Consider the effect of changing only the productivity level (ξ_b), holding other factors such as the pace of mortality decline unchanged. For the life-cycle model with disutility of labor specification, the marginal benefit of delaying retirement is monotonically decreasing in productivity level when $\sigma < 1$, but the marginal cost is independent of it. In an economic environment with both mortality decline and productivity increase (which is sometimes further assumed to have a constant growth rate), when the growth rate of productivity is above some threshold level (and thus, the productivity increase effect dominates), optimal age of retirement will be decreasing over cohorts eventually, as in Bloom et al. (2014).[12] On the other hand, for the life-cycle model with utility jump at retirement specification, we show in (8a) that optimal age of retirement is independent of productivity level even when $\sigma < 1$ because the productivity level affects marginal benefit and marginal cost equally.[13]

In the proposed definition of EDDR, we assume that the birth year is the only source of heterogeneity, as in the conventional dependency ratio and the forward-looking demographic definition. Under this assumption, individuals born at different years face two major differences: cohort-specific survival function as well as cohort-specific productivity level, both showing substantial changes during the last century. If the proposed EDDR is based on the model with the disutility of labor specification, then it implies that when either the growth rate of productivity is high enough or the extent of mortality decline gradually becomes less significant, the optimal age of retirement will decrease over cohorts. Thus, the

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

post-retirement dependency ratio will be increasing rapidly over time and in fact increasing even more rapidly than under the conventional dependency ratio. Since this prediction is implied by the disutility of labor specification, but not determined by the data, this may pose a problem in measuring resource dependency for advanced countries with increasing productivity and wealth. On the other hand, the utility jump at retirement specification does not suffer from this limitation. Because of this feature, we choose the utility jump at retirement specification in our proposed definition of EDDR.

3. QUANTITATIVE ANALYSIS: THE BASELINE CASE

Based on the theoretical model in the previous section, we now perform quantitative analysis. We aim to conduct an objective and transparent analysis. As far as possible, the specifications in our analysis are those commonly used in the literature.

There are demographic and economic elements in our proposed method. We first discuss the survival probabilities in Section 3.1. Based on the survival probabilities and the theoretical model in Section 2, we calculate cohort-specific optimal choices of S_b^* and R_b^* in Section 3.2. We then aggregate these cohort-specific choices to obtain the cross-sectional EDDRs from 2010 to 2060 in Section 3.3.[14]

3.1. Survival Probabilities of Different Cohorts

Since our objective is to construct the EDDR from 2010 to 2060, the above framework requires survival functions for every cohort from 1900 to 2060 continuously, under the assumption that the maximum biological age (Ω) is 110. On the other hand, we have only annual data of age-specific survival probabilities (from the Berkeley Mortality Database) from 1900 to 2000. To fill the missing data, we use the well-established model suggested by Lee and Carter (1992) to forecast the survival probabilities of different cohorts.[15]

In essence, the Lee–Carter approach is based on a log-additive model comprising a fixed age component and a time component to model the mortality rates in logarithmic scale. This parsimonious demographic model has been successively applied to the G7 and other countries to forecast life expectancy at birth; see Tuljapurkar et al. (2000) and Booth et al. (2006). A brief description of the Lee–Carter method is given in Appendix C.

Figure 2A presents plots of the observed and estimated values of survival functions for various US cohorts (males and females combined) in years 1900, 1950, 2000, and 2060. The solid lines denote the actual survival probabilities obtained from life tables, and the dotted lines estimates of the survival probabilities by the Lee–Carter model. As seen from these plots, the fitted curves are reasonably close to the actual observations, especially for cohorts after 1950. The survival distribution for year 1900 lies right at the bottom, with the other three distributions

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

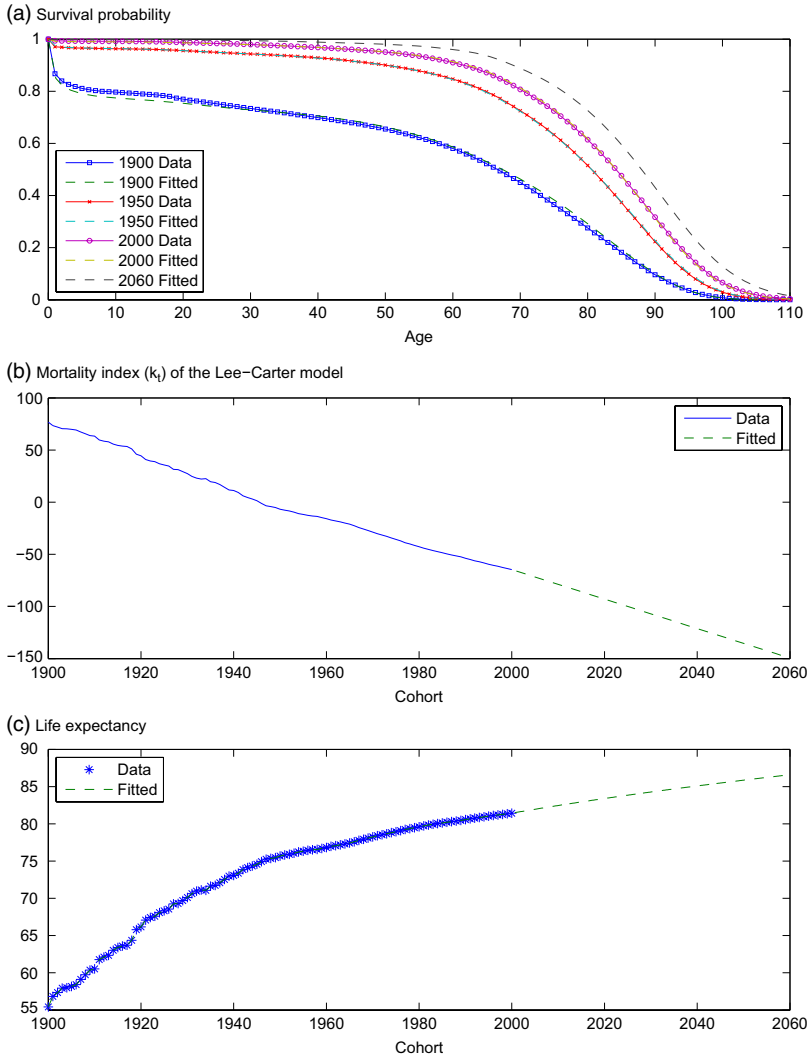


FIGURE 2. Demographic features of different cohorts.

shifting upward to the right side over each period of either 50 or 60 years. These upward shifts in the survival distributions imply substantial improvements in survival probability at all ages, with more prominent increments from 1900 to 1950 than those from 1950 to 2000. Such improvements are also captured by the steadily falling mortality level (k_t, defined in Appendix C) and rising life expectancy at birth over the period under study, which are displayed in Figure 2B and C, respectively.

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

3.2. Cohort-specific Schooling and Retirement Choices

To calculate S_b^* and R_b^* for different cohorts from 1900 to 2060 with fitted survival probabilities from the Lee–Carter method, we need to specify the human capital function, together with other parameter values of the model. For the human capital function, we assume that

$$h(S) = e^{\gamma S^{\lambda}}, \tag{18}$$

where $\gamma > 0$ and $0 < \lambda \leq 1$. This functional form is consistent with those in the literature, such as Cervellati and Sunde (2013) and Cai and Lau (2017). According to this specification, the rate of return to schooling is

$$\frac{h'(S)}{h(S)} = \gamma \lambda S^{\lambda - 1}, \tag{19}$$

which is a decreasing function in S when λ is strictly less than 1.

The parameter values in the baseline model are chosen to match those in the literature as far as possible. (We will also use other parameter values in the sensitivity analysis in Section 5.) In particular, we choose $\rho = 0.02$ and $r = 0.05$, following Barro et al. (1995). We also assume that $M = 10$, following Boucekkine et al. (2003).

The values of free parameters γ and λ in the human capital function (18), together with σ and θ, are chosen according to the following procedure. We use the non-linear least squares method to choose γ and λ, after imposing two restrictions: (a) the optimal age of retirement of the 1950 cohort is[16]

$$R_{1950}^* = 66 - M, \tag{20}$$

and (b) the percentage drop of consumption at retirement is 15%.[17] According to (5), the latter condition implies

$$(1+\theta)^{\sigma - 1} = 1 - \%_\text{drop_consumption} = 0.85. \tag{21}$$

Under these two restrictions and other parameter values, we choose γ and λ to minimize the sum of squared residuals (SSR)

$$SSR_S(\gamma, \lambda) = \sum_{b=1900}^{1975} \left\{ S_b^*(\gamma, \lambda; \sigma(\gamma, \lambda), \theta(\gamma, \lambda)) - \left[S_b^{\text{data}} - (M-6) \right] \right\}^2, \tag{22}$$

where $S_b^*(\gamma, \lambda; \sigma(\gamma, \lambda), \theta(\gamma, \lambda))$ is the optimal years of schooling of cohort-b individuals calculated from the model, and S_b^{data} is the years of schooling data from Goldin and Katz (2008).[18] We focus on the 76 observations, from 1900 to 1975, covered in both the Berkeley Mortality Database and Goldin and Katz (2008). Specifically, our search method allows that γ ranges from 0 to 10, and λ ranges from 0 to 1. For each pairs of (γ, λ) in these selected ranges, we first use (7), (8a), (20), (21), and the fitted survival probabilities at 1950 to calculate $\sigma(\gamma, \lambda)$ and $\theta(\gamma, \lambda)$. We then use (7) and (8a) to obtain the paths of $R_b^*(\gamma, \lambda; \sigma(\gamma, \lambda), \theta(\gamma, \lambda))$

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

TABLE 1. Parameter calibration

Parameters	Description	Value	Source/method
ρ	Subjective discount rate	2%	Barro et al. (1995)
r	Real interest rate	5%	Barro et al. (1995)
M	The age that individuals begin making economic decisions	10	Boucekkine et al. (2003)
Ω	Maximum age in the model	110	Our assumption (see footnote 14)
γ	Free parameter in human capital function	0.076	SSR minimization subject to (20), (21)[a]
λ	Free parameter in human capital function	0.913	SSR minimization subject to (20), (21)
σ	Intertemporal elasticity of substitution	0.794	SSR minimization subject to (20), (21)
θ	Utility gain from leisure during retirement	1.199	SSR minimization subject to (20), (21)

[a]SSR minimization subject to (20): $R^*_{1950} = 66 - M$ and (21): 15% drop in consumption at retirement.

and $S_b^* (\gamma, \lambda; \sigma (\gamma, \lambda), \theta (\gamma, \lambda))$. The corresponding values of SSR and root mean squared error (RMSE) according to (22) can also be obtained.

Using the above procedure, the least-squares estimates are found to be $\widehat{\gamma} = 0.076$ and $\widehat{\lambda} = 0.913$. The corresponding values of $\sigma (\widehat{\gamma}, \widehat{\lambda})$ and $\theta (\widehat{\gamma}, \widehat{\lambda})$ are 0.794 and 1.199, respectively. The parameter values of the baseline case are summarized in Table 1. In particular, it is noted that the estimated value of the curvature parameter (λ) in human capital function is 0.913, with the rate of return to schooling mildly decreasing over the relevant range.

The path of years of schooling for different cohorts from 1900 to 2040 and the path of age of retirement from 1900 to 1990 are plotted in Figure 3A and B, respectively. In particular, it is observed that over the years from 1900 to 1975, the fitted years of schooling, which are obtained from minimizing the SSR according to equation (22), match reasonably well with the actual years of schooling data by Goldin and Katz (2008). The computed RMSE is 0.196 years, which is relatively small by the conventional standard.

As observed in Figure 3, both years of schooling and age of retirement increase during the whole period under study. Moreover, the rate of change slows down from the post-WW2 period onwards. For age of retirement, it increases by 7.17 years (from 56.92 to 64.09) from the 1900 cohort to the 1940 cohort, but it only increases by 4.62 years (from 66 to 70.62) from the 1950 cohort to the 1990 cohort.[19] Similarly, the optimal years of schooling increase by 3.17 years (from 9.05 to 12.22) from the 1900 cohort to the 1940 cohort. The trend growth rate slows down, with years of schooling increasing by 2.1 years (from 13.09 to 15.19) from the 1950 cohort to the 1990 cohort and by 2.02 years (from 15.63 to 17.65)

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

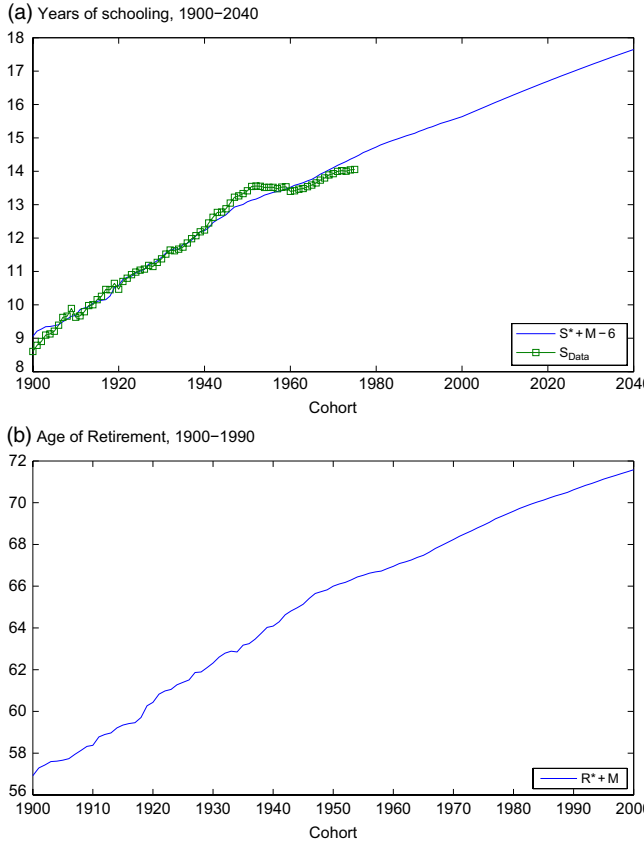


FIGURE 3. Optimal schooling and retirement decisions of different cohorts.

from the 2000 cohort to the 2040 cohort. In summary, our computational analysis using the US data indicates that years of schooling and age of retirement move in the same direction in response to mortality decline, and the magnitude of the changes slows down after WW2 during which the magnitude of mortality decline also diminished.

3.3. Total Dependency Ratio and Its Components: Cross-sectional Summation of Cohort-specific Choices

In order to compute the pre-working, post-retirement, and total dependency ratios, we need to use the population data at a given year t, which consists of all surviving persons from the cohort starting at year $t - \Omega$ to the cohort ending at year t. Figure 4 displays the annual births in the USA from 1900 to 2060. There was an acute and sustained increase in annual births from the 1940s into the 1950s, which ended in the mid-1960s. This is the so-called period of baby boomers in the

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

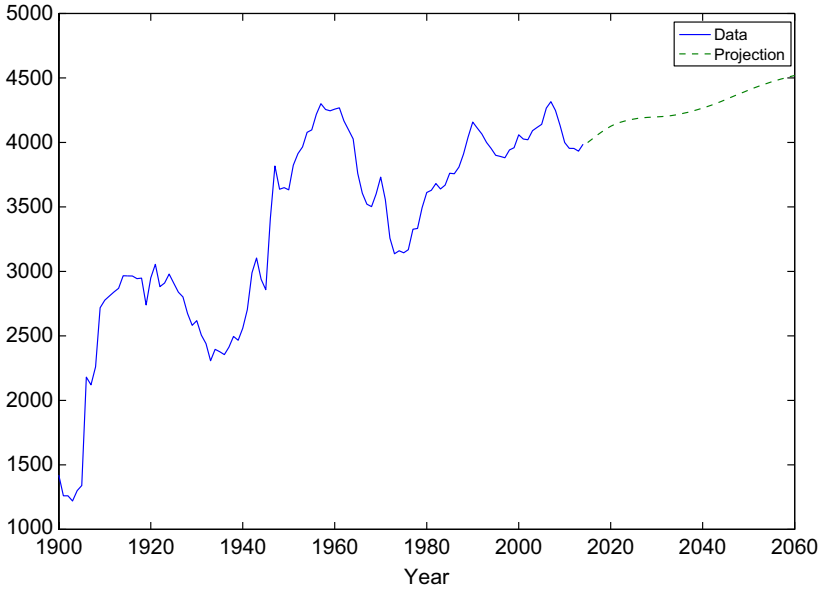


FIGURE 4. US births number (in thousands).

USA between 1946 and 1964, with annual births consistently above 3.4 million. Since then the births declined to a low of about 3.1 million in the mid-1970s. The annual number of births increased again into the 1980s and 1990s but with a smaller magnitude of the derivation above the long-term trend.

We now turn to the computation of pre-working and post-retirement dependency ratios as specified in equations (15) and (16). We conduct the exercise under the assumption of no immigration. The common denominator of these ratios represents the working-age population at year t. It consists of the sum of surviving workers with ages right after the optimal years of schooling and before the optimal age of retirement, whom we can identify by checking the range of the cohort-varying years of schooling and the cohort-varying age of retirement for each cohort stretching from year $t - \Omega$ to year t. The number of pre-working survivors at year t consists of those who are at ages of pre-schooling and schooling for each cohort. Similarly, the number of surviving persons in the post-retirement era comprises those who are beyond the optimal retirement ages over the cohorts under study. Plots of the *EDDR* and its components from 2010 to 2060 are displayed in Figure 5A; the underlying pre-working, post-retirement, and working-age populations are displayed in Figure 5B.

As can be gleaned from these plots, the post-retirement dependency ratio starts from 0.23 in 2010, increasing steadily to 0.31 in 2033 and then declining slowly to 0.24 in 2060. The peak of 0.31 around year 2033 can be explained by the effect of baby boom cohorts on the age structure of the US population. As the baby boomers of 1946 cohort begin to turn to retirement age, which is around

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

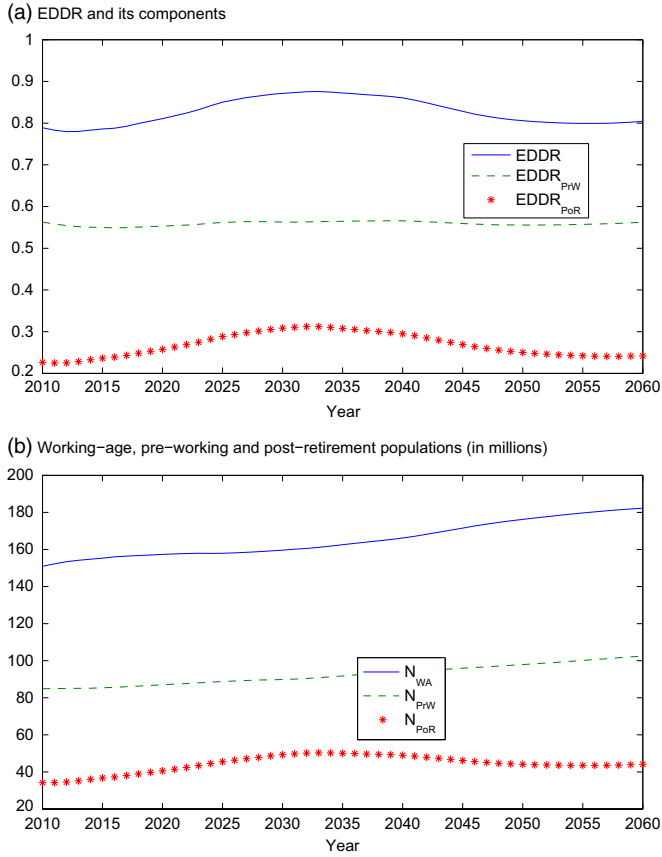


FIGURE 5. Economic-demographic dependency ratios.

65 according to Figure 3B, it increases the share of the population in the post-retirement group in year 2011. Such an effect will continue for about 20 years, reaching the peak of 0.31 by 2033, and then declining to 0.25 in 2050 when most of the baby boomers reach the oldest-old group.[20] On the other hand, the baby boom cohorts mainly affect the size of the pre-working group up to mid-1980s, and thus, the path of the pre-working dependency ratio fluctuates less for the period of 2010 to 2060. It hovers around 0.56 throughout the entire period from 2010 to 2060, with a minimum of 0.55 in 2015 and a maximum of 0.57 in 2039.

The path of the *EDDR*, which is simply the sum of the pre-working and post-retirement dependency ratios, reflects the dependency of those in the pre-working and post-retirement groups relative to the size of the working-age population across cohorts. As can be seen from Figure 5A, the total dependency ratio begins with 0.79 in year 2010, increasing slowly until reaching its peak at 0.87 in 2033. It then declines steadily to 0.80 in 2060. We observe that the shape of the total

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

dependency ratio is very similar to that of the post-retirement dependency ratio. This is not surprising as the path of pre-working dependency ratio does not change much. As a result, the long-term influence of the baby boomers on the post-retirement dependency ratio, through the age structure of the US population, is also reflected in the path of the total dependency ratio, in an almost one-to-one fashion.

4. COMPARISON WITH THE CONVENTIONAL DEFINITION

In this section, we compare the EDDRs with the conventional dependency ratios, assuming no immigration in both cases. The conventional definition of the child dependency ratio is given by

$$DR_C(t) = \frac{N_C(t)}{N_A(t)} = \frac{\int_{t-20}^{t} B(b)\,\widetilde{l}_b(t-b)\,db}{\int_{t-65}^{t-20} B(b)\,\widetilde{l}_b(t-b)\,db}, \tag{23}$$

and the old-age dependency ratio is given by

$$DR_{OA}(t) = \frac{N_{OA}(t)}{N_A(t)} = \frac{\int_{t-\Omega}^{t-65} B(b)\,\widetilde{l}_b(t-b)\,db}{\int_{t-65}^{t-20} B(b)\,\widetilde{l}_b(t-b)\,db}, \tag{24}$$

where $N_C(t)$, $N_A(t)$, and $N_{OA}(t)$ represent, respectively, the number of child, adult, and old-age population at year t, and other terms follow the same definitions as above. For a given year t, the child and the old-age dependency ratios are calculated according to the demographic distribution of all cohorts from year $t-\Omega$ to year t, with the adult population being all those who are between 20 and 65, the child population being all those below 20, and the old-age population being all those beyond 65. The terms DR_C and DR_{OA} in equations (23) and (24) are similar, but not identical, to $EDDR_{PrW}$ and $EDDR_{PoR}$ in (15) and (16).

We now investigate any under- or over-projection of the conventional dependency ratios from 2010 to 2060, as compared with the behaviorally based EDDR. First, we observe in Figure 6 that the conventional child dependency ratio (DR_C) under-projects $EDDR_{PrW}$ throughout the interval, with 4.5 percentage points in 2010 and then gradually increasing to about 6 to 7 percentage points after 2035. Second, the conventional old-age dependency ratio (DR_{OA}) is very close to but slightly lower than $EDDR_{PoR}$ in 2010, but starts to over-project over this period, with the extent of over-projection increasing to 11 percentage points in 2060. Combining these two results, the conventional total dependency ratio under-projects by about 5.0 percentage points in 2010, with the under-projection gap narrowing down steadily to zero in 2045 and over-projecting afterwards by about 4.0 percentage points in 2060.[21]

Besides the above descriptive summary, a more in-depth comparison of the conventional and our proposed measures of dependency ratio, as well as an understanding of the underlying reasons of the above differences, is not straightforward. There are both short- and long-term changes in the conventional dependency

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

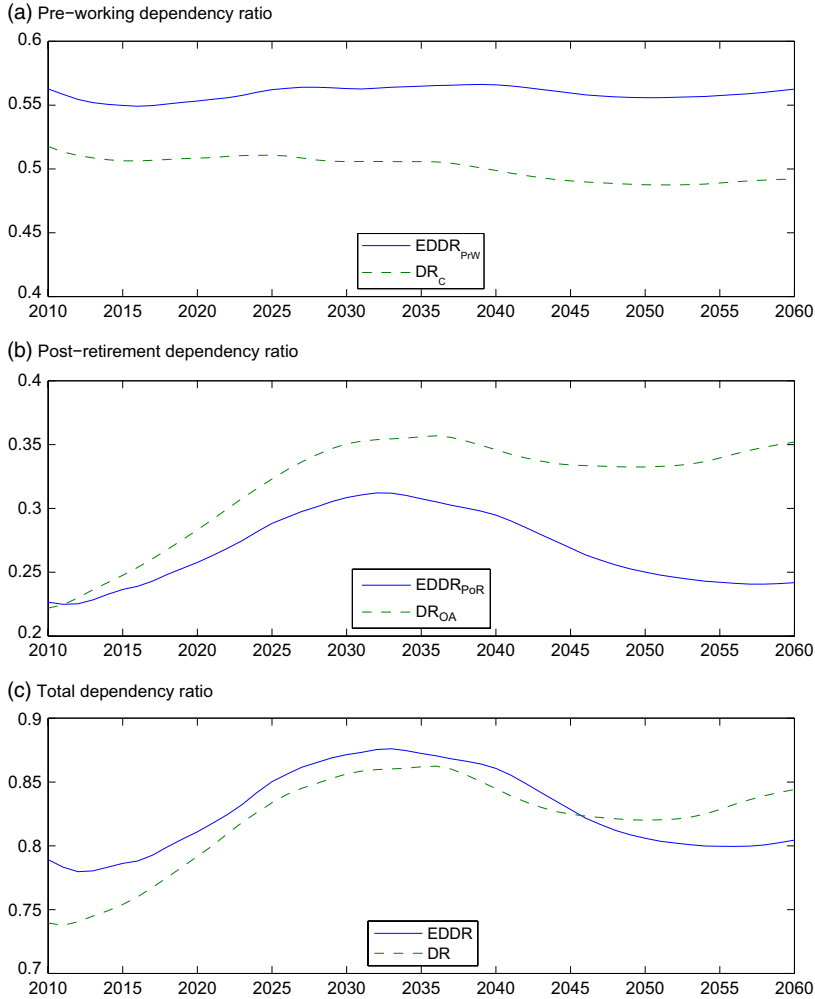


FIGURE 6. EDDR and the conventional definition.

ratios and EDDRs, as discussed earlier in this section and in Section 3.3. Moreover, it is observed in Figure 6 that the initial levels (at year 2010) according to the two measures are different. Since our primary purpose is to study *the effect on resource dependency of rapid population aging in the coming decades*, we propose to (a) decompose the effect under investigation into two parts: due to initial level and due to the change from the initial period of 2010 to a subsequent period,[22] (b) downplay the difference in initial levels and focus mainly on the path of average *growth rate* of various component terms of the dependency ratio from 2010 to a subsequent year,[23] and (c) downplay the short-term movement, which is more likely to be affected by the size of the baby boom cohorts (1946–1964)

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

and that of their children (the so-called echo boomers) and examine mainly the *longer term effect* on the later period of the sample, say from 2040 to 2060.

We find it useful to use the method of expanding duration, which is similar to the idea of recursive (expanding window) estimates and forecasts in time series analysis [see, e.g., Brown et al. (1975), Clark and West (2007), and Rapach et al. (2008)], to calculate the annualized growth rate of the component terms of a dependency ratio. Specifically, for each expanding duration starting from year 2010 to a future year t (which ranges from 2011 to 2060), we can compute the annualized growth rate (i.e., average growth rate) over a duration of $t - 2010$ years.[24] As such, a path consisting of 50 average growth rates over the corresponding expanding durations can be computed accordingly. Such a path of recursive annualized growth rates provides information about their stability through the period of study.

As displayed in Figure 7A, the path of average growth rates of child population increases rapidly from –0.17% in 2011 to approximately 0.11% in 2025, thereafter growing very slowly to 0.18% in 2060. The path of average growth rates of adult population decreases rapidly from 0.67% in 2011 to 0.18% in 2025, thereafter bouncing back slightly to reach 0.28% in 2060. For the hump-shaped path of average growth rates of old-age population, it begins with 1.84% in 2011, increasing rapidly to 2.61% in 2014, and remaining relatively unchanged until 2028, thereafter declining gradually to 1.21% in 2060.[25]

The paths of the average growth rates of component terms under the proposed approach can be observed in Figure 7B. Except for the first few years, the path of average growth rates of pre-working population is consistently flat and small in magnitude. It grows rapidly from 0.15% in 2011 to 0.30% in 2025, thereafter increasing relatively slowly to 0.38% in 2060. The path of average growth rates of working-age population decreases rapidly from 0.90% in 2011 to 0.30% in 2025, thereafter bouncing back gradually to reach 0.38% in 2060. Regarding the average growth rates of post-retirement population, it starts from 0.16% in 2011, increasing substantially to 1.93% in 2025, and then declining moderately to about 0.51% in 2060.

We find it helpful to examine the growth rates of various ratios of the conventional and new definitions, given by (15), (16), (23), and (24). It is observed from Figure 7C that the growth rate of child dependency ratio, which is the difference of the growth rate of child population and that of the adult population, is always negative. Similarly, it is seen from Figure 7D that the growth rate of old-age dependency ratio is always positive. Therefore, the conventional child dependency ratio (resp. conventional old-age dependency ratio) in the future will be lower (resp. higher) than the initial level in year 2010. Moreover, the annualized growth rate (in absolute magnitude) of the old-age population is above 1% most of the time, which is much higher than that of the child population. The intuition of these patterns is that the age-specific mortality rates for children until around age 20 are already very low for advanced countries such as USA, and there is not much room for further improvement. Continuous mortality decline in the coming

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

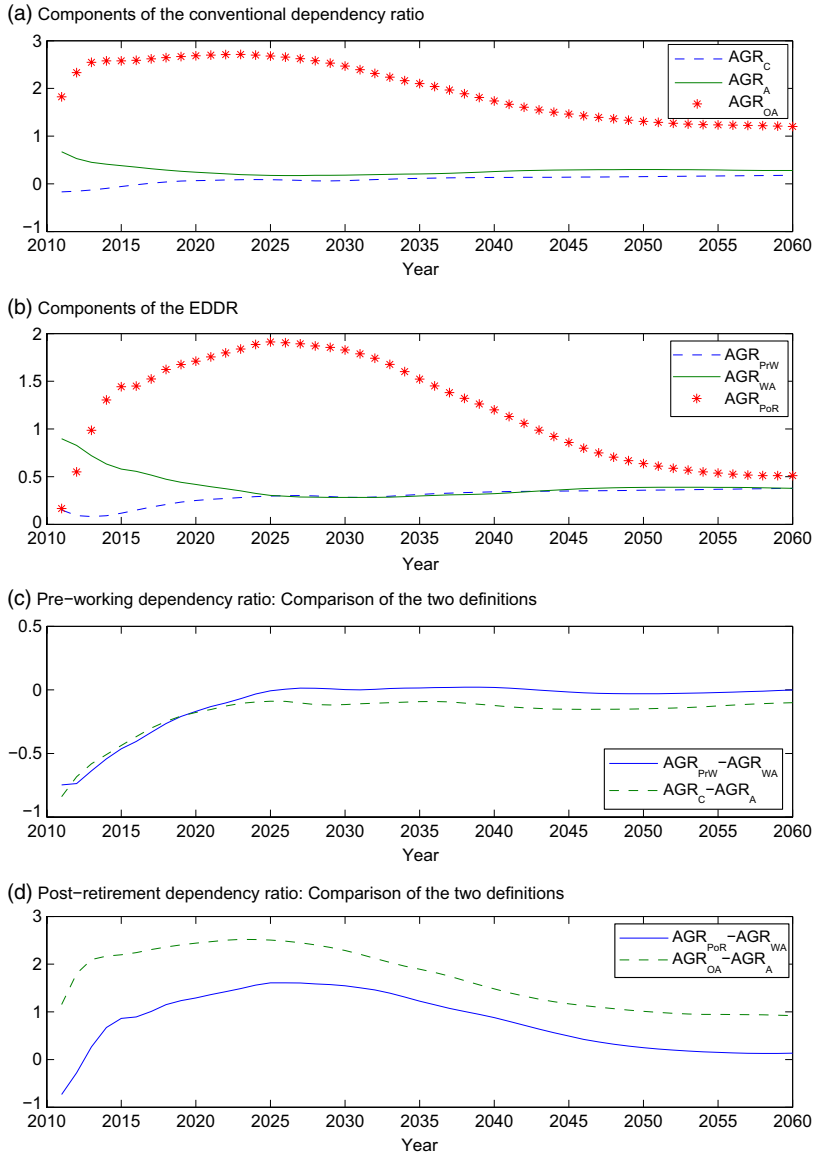


FIGURE 7. Annualized growth rate (in percentage) of various terms.

decades will come from adults and especially in old age. As a result of these demographic changes, adult population will grow faster than child population (see Figure 7C), and old-age population grows even at a faster rate (see Figure 7D). When the cutoff points of child and old age are unchanged, as in the conventional approach, the child dependency ratio will decrease but the old-age dependency

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

ratio will increase more drastically in the coming decades. Combining the mildly declining child dependency ratio (Figure 6A) and the strongly increasing old-age dependency ratio (Figure 6B), it is concluded that the total dependency ratio will increase quite significantly (Figure 6C) according to the conventional definition.

We now compare the annualized growth rate of pre-working (or child) and post-retirement (or old-age) dependency ratios under the two approaches to investigate any over- or under-projection. As observed in Figure 7C, the growth rate of pre-working dependency ratio of our proposed approach is negative from 2011 to around 2025 and is around 0% afterwards. As a result, after 2025 when the effect due to fluctuating birth numbers from 1980s to 2000s becomes less important, there is no strong upward nor downward trend in $EDDR_{PrW}$ in the coming decades (see Figure 6A also), and the growth rate of pre-working dependency ratio is always higher than that of child dependency ratio, which is negative (Figure 7C). Compared with the proposed EDDR, *conventional child dependency ratio under-projects from 2025 to 2060*. On the other hand, it is observed in Figure 7D that the shape of the growth rate of post-retirement dependency ratio against year is somewhat similar to that of the conventional old-age dependency ratio, except that the growth rate of $EDDR_{PoR}$ is much lower (e.g., 0.13% in 2060 versus 0.93% in 2060) than that of the conventional old-age dependency ratio. Thus, the overall increase in $EDDR_{PoR}$ from 2010 to 2060 is much smaller than that in DR_{OA} (Figure 6B).[26] Compared with the proposed EDDR, *conventional old-age dependency ratio always over-projects from 2010 to 2060*. Moreover, the difference of the two growth rate lines in Figure 7D is relatively large when compared with that in Figure 7C, implying that the magnitude of *over-projection of the conventional old-age dependency ratio dominates* that of under-projection of the conventional child dependency ratio.

In summary, the conventional total dependency ratio increases by 0.105 from 2010 to 2060, but the EDDR increases only by 0.015.[27] Relative to the EDDR, the conventional dependency ratio over-projects the effect of resource dependence due to population aging in these five decades by 0.09 (i.e., 86% of 0.105).

5. OVER-PROJECTION OF THE CONVENTIONAL DEPENDENCY RATIO: SENSITIVITY ANALYSIS

In this section, we examine whether the result that the conventional dependency ratio over-projects the effect of population aging from 2010 to 2060 by 0.09 is sensitive or not to changes in the parameters of the baseline model.

We consider 10 cases. In each case, we allow for deviation in one parameter from its value specified in the baseline model while keeping values of all other parameters intact. The first sensitivity analysis we perform is to change the value of real interest rate from $r = 5\%$ to $r = 4\%$ (Case 1), and to $r = 6\%$ (Case 2). Second, we examine how our main results are sensitive with respect to the subjective discount rate when we change from $\rho = 2\%$ to $\rho = 1.5\%$ (Case 3) and to $\rho = 2.5\%$ (Case 4). Third, we consider different values of the age that agents

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

begin making economic decisions. Instead of assuming $M = 10$, we analyze two cases: $M = 8$ (Case 5) and $M = 12$ (Case 6). Fourth, we change the percentage drop of consumption at retirement from 0.15 to 0.10 (Case 7) and to 0.20 (Case 8). Finally, we conduct analysis on calibrating different ages of retirement according to the chosen cohort years. Instead of calibrating the age of retirement to the 1950 cohort as $R^*_{1950} = 66 - M$, we calibrate the age of retirement to the 1945 cohort as $R^*_{1945} = 66 - M$ (Case 9) and to the 1955 cohort as $R^*_{1955} = 66.17 - M$ (Case 10).[28]

Table 2 tabulates quantitative results for the 10 cases of sensitivity analysis. Column 1 contains estimates of free parameters γ, λ, θ, and σ. It is observed that their magnitudes are relatively close to the baseline case. In particular, λ ranges quite narrowly from 0.882 to 0.933, and σ ranges from 0.738 to 0.855, which are similar in magnitude to the values estimated by many researchers.[29]

Column 2 reports RMSE for each of the 10 cases. It ranges from 0.176 to 0.246, indicating that the schooling years series computed are all reasonably close to the schooling data in Goldin and Katz (2008). Columns 4 and 5 give estimates of the optimal years of schooling and age of retirement, respectively, of four selected cohorts, including 1900 in the beginning of the period under consideration, 1950 in the middle, and 1990 (with age of retirement only) and 2040 (with years of schooling only) toward the end. It is observed that the optimal years of schooling (ranging from 17.24 to 18.40 in 2040) and the optimal age of retirement (ranging from 70.20 to 71.50 in 1990) are quite robust to parameter changes which are within reasonable ranges.[30]

We now turn to the dependency ratios. Column 7 of Table 2 reports the pre-retirement dependency ratios ($EDDR_{PrW}$) of the proposed method at years 2010 and 2060, respectively, for the 10 cases of sensitivity analysis, and Column 8 reports the corresponding post-retirement dependency ratios ($EDDR_{PoR}$). The time paths of $EDDR_{PrW}$ and $EDDR_{PoR}$ are presented in Figure 8A and B. Generally speaking, there are two major observations for each series. First, in all cases except Case 6 ($M = 12$), the path of $EDDR_{PrW}$ is relatively flat, hovering around 0.56 over the period 2010–2060. Second, the path of $EDDR_{PrW}$ for Case 6 looks somewhat like a "parallel" upward shift. To isolate the effect of the initial level and speed of change of population aging during this period, we examine Figure 9A. We observe that the paths of the annualized growth rates of $EDDR_{PrW}$ are very close to 0 for years beyond 2025.[31]

A similar, but slightly more diverse, pattern is observed for the post-retirement dependency ratios. First, in all cases except Case 9 ($R^*_{1945} = 66 - M$), there is a hump shape in $EDDR_{PoR}$, starting around 0.22 at year 2010, rising to around 0.30 in early 2030s, and then going down to somewhere between 0.23 and 0.25 at year 2060. Second, the path of $EDDR_{PoR}$ for Case 9 somewhat resembles a "parallel" downward shift. According to Figure 9B, the paths of the annualized growth rates of $EDDR_{PoR}$ are between 0.03% and 0.21% in year 2060. The range of variation of the annualized growth rates of $EDDR_{PoR}$ is larger than that of $EDDR_{PrW}$ (from –0.03% to 0.03%).

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

TABLE 2. Sensitivity Analysis

Specification	RMSE	Cohort-specific choices				Dependency ratios				
Baseline	0.196		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.076$		1900	9.05	56.92	2010	0.5624	0.2265	0.5176	0.2219	–0.049
$\lambda = 0.913$		1950	13.09	66	2060	0.5623	0.2420	0.4921	0.3520	0.040
$\theta = 1.199$		1990		70.62	Diff.	–0.0001	0.0155	–0.0255	0.1301	**0.0892**
$\sigma = 0.794$		2040	17.65		AGR	**0%**	**0.13%**	**–0.10%**	**0.92%**	
$r = 4\%$	0.211		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.072$		1900	9.11	56.31	2010	0.5663	0.2284	0.5176	0.2219	
$\lambda = 0.882$		1950	13.07	66	2060	0.5665	0.2313	0.4921	0.3520	
$\theta = 1.681$		1990		71.28	Diff.	0.0002	0.0029			0.1015
$\sigma = 0.835$		2040	17.94		AGR	0%	*0.03%*			
$r = 6\%$	0.187		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.082$		1900	9.01	57.17	2010	0.5600	0.2255	0.5176	0.2219	
$\lambda = 0.932$		1950	13.10	66	2060	0.5581	0.2473	0.4921	0.3520	
$\theta = 0.905$		1990		70.30	Diff.	–0.0019	0.0218			0.0847
$\sigma = 0.748$		2040	17.43		AGR	–0.01%	0.18%			
$\rho = 1.5\%$	0.195		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.074$		1900	9.04	57.66	2010	0.5618	0.2252	0.5176	0.2219	
$\lambda = 0.920$		1950	13.09	66	2060	0.5671	0.2505	0.4921	0.3520	
$\theta = 1.040$		1990		70.20	Diff.	0.0053	0.0253			0.0740
$\sigma = 0.772$		2040	17.68		AGR	0.02%	*0.21%*			

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

TABLE 2. Continued

Specification	RMSE	Cohort-specific choices				Dependency ratios				
$\rho = 2.5\%$	0.196		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.078$		1900	9.06	55.91	2010	0.5627	0.2281	0.5176	0.2219	
$\lambda = 0.902$		1950	13.09	66	2060	0.5543	0.2315	0.4921	0.3520	
$\theta = 1.414$		1990		71.15	Diff.	–0.0084	0.0034			0.1096
$\sigma = 0.816$		2040	17.54		AGR	*–0.03%*	*0.03%*			
$M = 8$	0.176		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.082$		1900	8.94	56.90	2010	0.5571	0.2250	0.5176	0.2219	
$\lambda = 0.893$		1950	13.11	66	2060	0.5527	0.2481	0.4921	0.3520	
$\theta = 1.300$		1990		70.21	Diff.	–0.0044	0.0231			0.0859
$\sigma = 0.805$		2040	17.24		AGR	–0.02%	0.20%			
$M = 12$	0.246		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.069$		1900	9.25	57.16	2010	0.5716	0.2288	0.5176	0.2219	
$\lambda = 0.933$		1950	13.04	66	2060	0.5809	0.2327	0.4921	0.3520	
$\theta = 1.105$		1990		71.32	Diff.	0.0093	0.0039			0.0914
$\sigma = 0.782$		2040	18.40		AGR	*0.03%*	0.03%			
C_drop= 0.1	0.196		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.074$		1900	9.05	57.42	2010	0.5623	0.2257	0.5176	0.2219	
$\lambda = 0.918$		1950	13.09	66	2060	0.5662	0.2475	0.4921	0.3520	
$\theta = 1.068$		1990		70.35	Diff.	0.0039	0.0218			0.0789
$\sigma = 0.855$		2040	17.70		AGR	0.01%	0.18%			

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

TABLE 2. Continued

Specification	**RMSE**	**Cohort-specific choices**				**Dependency ratios**				
C_drop = 0.2	0.196		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.077$		1900	9.06	56.38	2010	0.5624	0.2273	0.5176	0.2219	
$\lambda = 0.907$		1950	13.09	66	2060	0.5578	0.2366	0.4921	0.3520	
$\theta = 1.347$		1990		70.89	Diff.	–0.0046	0.0093			0.0999
$\sigma = 0.738$		2040	17.58		AGR	–0.02%	0.08%			
$R^*_{1945} = 66 - M$	0.194		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.075$		1900	9.04	57.69	2010	0.5561	0.2132	0.5176	0.2219	
$\lambda = 0.915$		1950	13.09	66.87	2060	0.5542	0.2253	0.4921	0.3520	
$\theta = 1.166$		1990		71.50	Diff.	–0.0019	0.0121			0.0944
$\sigma = 0.790$		2040	17.61		AGR	–0.01%	0.11%			
$R^*_{1955} = 66.17 - M$	0.197		S*+M-6	R*+M		$EDDR_{PrW}$	$EDDR_{PoR}$	DR_C	DR_{OA}	$DR - EDDR$
$\gamma = 0.076$		1900	9.06	56.60	2010	0.5652	0.2318	0.5176	0.2219	
$\lambda = 0.912$		1950	13.09	65.65	2060	0.5662	0.2491	0.4921	0.3520	
$\theta = 1.213$		1990		70.27	Diff.	0.001	0.0173			0.0863
$\sigma = 0.795$		2040	17.67		AGR	0%	0.14%			

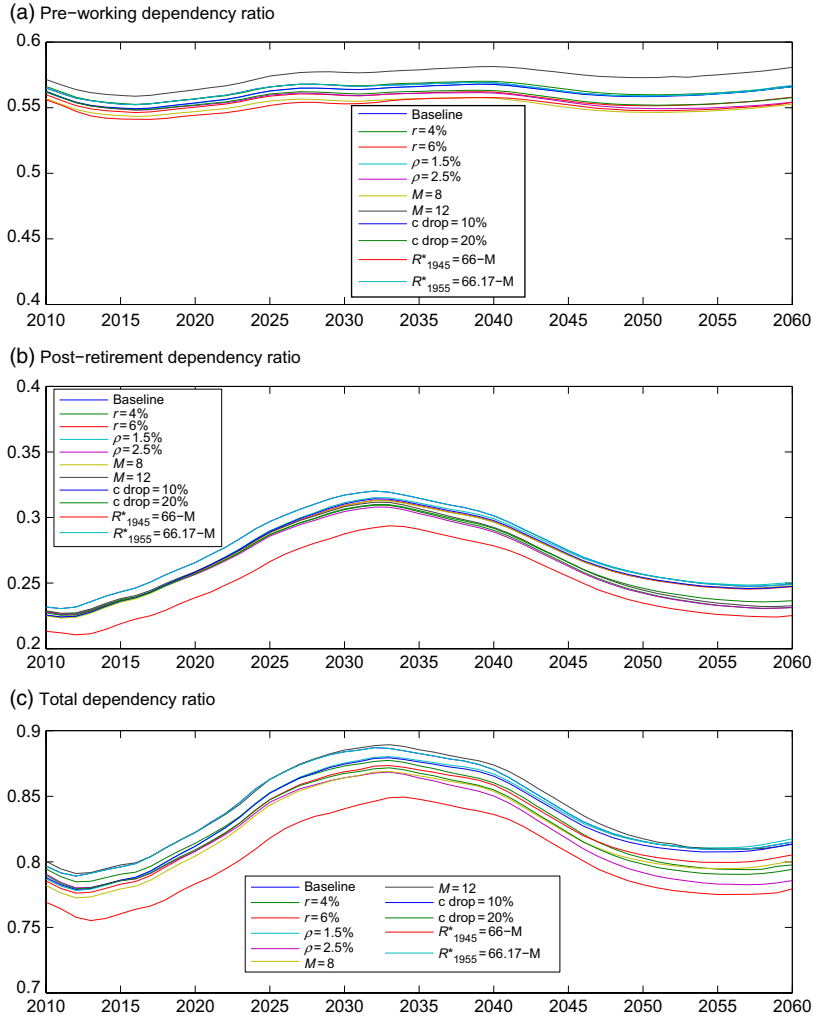


FIGURE 8. Sensitivity analysis of EDDR.

Finally, we combine the above results to examine the total dependency ratio under the proposed approach. The entire paths of *EDDR* for all 10 cases over the period 2010–2060 are displayed in Figure 8C. It is observed that 9 out of the 10 paths of the total dependency ratios are close to that of the baseline case. The remaining path at the bottom of Figure 8C corresponds to Case 9 with $R^*_{1950} = 66 - M$, which is still reasonably close to the baseline path.

Columns 9 and 10 of Table 2 tabulate the dependency ratios of child (DR_C) and old-age (DR_{OA}) at years 2010 and 2060 and the difference of dependency ratios over the period 2010–2060. Column 11 reports the difference of the total

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

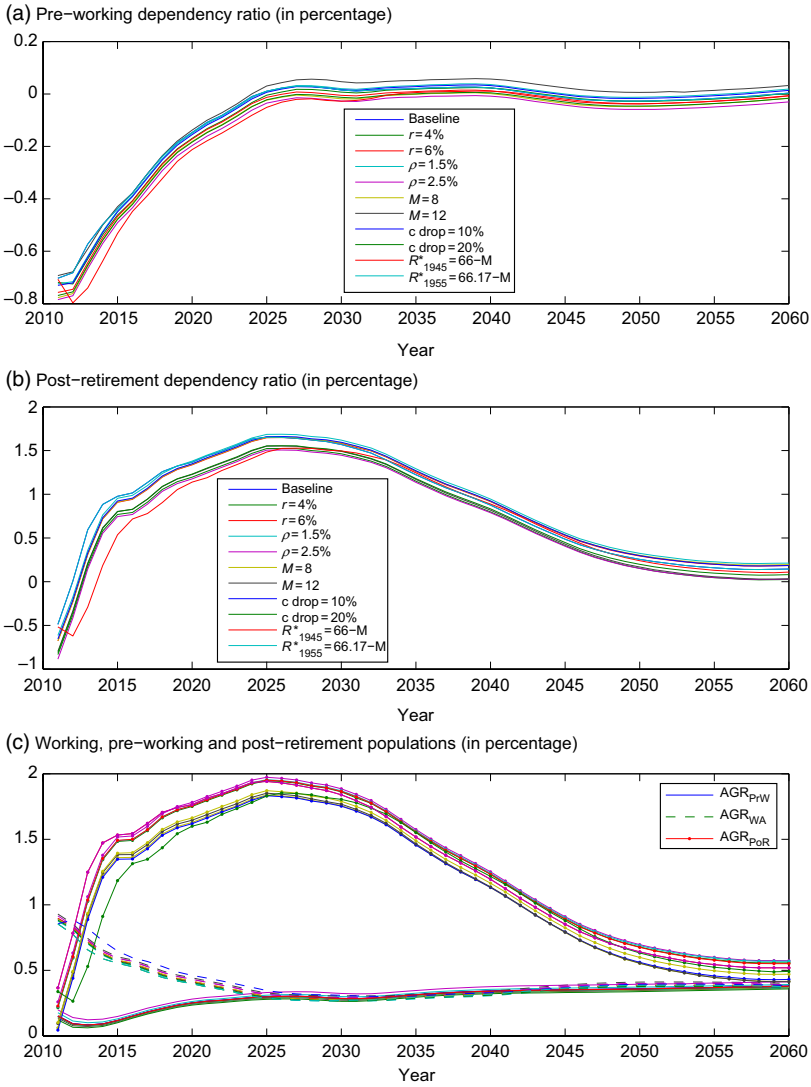


FIGURE 9. Sensitivity analysis of annualized growth rates.

dependency ratio under the conventional approach (*DR*) and the total dependency ratio under the economic-demographic approach (*EDDR*). For the baseline case, when compared with the projected increase by 0.015 in *EDDR* over the period 2010–2060, the conventional total dependency ratio over-projects the effect of population aging over the five decades by 0.09. For the 10 cases of sensitivity analysis, the change of $DR - EDDR$ from 2010 to 2060 ranges from 0.074 to 0.110, with the lower bound of this difference corresponding to Case 3 ($\rho = 1.5\%$)

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

and the upper bound corresponding to Case 4 ($\rho = 2.5\%$). Our findings indicate that the difference between *DR* and *EDDR* is quite sensitive to the subjective discount rate (ρ), but not to other parameters. When we ignore the two sensitive cases involving changes in ρ and only consider all other cases of parameter changes, the change of $DR - EDDR$ from 2010 to 2060 ranges from 0.079 to 0.102.[32] Based on these eight cases, we conclude that 75% to 97% of the 10.5 percentage point increase of the conventional dependency ratio in these five decades is likely to be caused by over-projection because the conventional approach does not account for behavioral response to demographic changes.[33]

6. CONCLUSION

Facing demographic changes, people may respond by adjusting their schooling and retirement behavior.[34] The conventional definition of dependency ratio, which is based on cohort-invariant cutoff points and does not take into account these behavioral responses, is likely to measure the impact of population aging incorrectly, especially when its pace is rapid.

To improve on the conventional dependency ratio, we propose a behaviorally based definition of dependency ratio, such that the cutoff points are based on a well-articulated economic model and may be cohort varying.[35] On the other hand, we intend to keep other aspects of the conventional definition unchanged in our modified definition. We believe that this is a good strategy because of two reasons: (a) the conventional dependency ratio is transparent for projection, and this good quality should be kept, and (b) any different conclusions based on these two definitions will be due to cohort-invariant versus cohort-varying cutoff points only. To achieve this objective, we use a life-cycle model focusing on the extensive margin of labor supply (versus schooling when young, and versus retirement when old) and use the utility jump at retirement specification such that optimal schooling years and age of retirement are independent of productivity change.[36] We maintain an essential feature of the demographic approach that birth year is the only source of heterogeneity. For easy comparison, we apply the proposed method to an environment without immigration.

We provide quantitative analysis of cohort-specific schooling and retirement choices based on the behavioral model and then obtain the EDDR by cross-sectional summation. To be as neutral as possible, we choose functional forms and parameter values commonly used in the literature. We then determine whether the conventional dependency ratio over- or under-projects, when compared with the behaviorally based definition. We find that in the baseline model, out of an increase of conventional dependency ratio of 0.105 from 2010 to 2060, 0.09 (or 86%) is due to over-projection. Our analysis suggests that to the extent that people adjust schooling and retirement decisions in response to mortality decline, conventional dependency ratio which does not account for these responses may not measure accurately the impact of population aging in the coming decades. Our sensitivity analysis suggests that while the results are a bit sensitive to changes

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

in one parameter (subjective discount rate ρ), they are relatively insensitive with respect to all other four parameters. Based on the eight cases (ignoring changes in ρ), a high proportion (75% to 97%) of the increase of 0.105 in conventional dependency ratio from 2010 to 2060 is due to over-projection.[37]

While we think the proposed economic model provides a useful framework to examine whether or not the conventional dependency ratio measures the impact of rapid population aging accurately, two natural questions arise. They are (a) how is our approach compared with others that are also based on economic concepts, such as the economic dependency ratio and NTA support ratio and (b) whether the proposed model is the most appropriate one in describing people's behavior in this context. A comparison reveals that both the economic dependency ratio and the EDDR focus on dependency in terms of labor resources, while the NTA support ratio is defined more broadly, including transfer of financial resources from family members or the government. Moreover, the economic dependency ratio includes both intensive and extensive margins of labor supply, while our definition focuses only on the extensive margin. One advantage of our proposed definition is that by focusing on a key difference from the conventional dependency ratio, it offers the unique advantage of singling out the role of cohort-invariant cutoff points of the conventional dependency ratio in determining the magnitude of over-projection of the impact of population aging. To the best of our knowledge, this paper is the first to document the discrepancy between dependency ratios based on cohort-invariant and cohort-varying cutoff points. Another advantage is that we provide a simple modification of a standard life-cycle model. The underlying economic reasons are clear and transparent, and the results can easily be replicated. On the other hand, various economic factors are captured in the economic dependency ratio or NTA support ratio. As a result, it is harder to isolate the separate contributions of each one of the several factors involved. Generally, the various alternative definitions are complementary, with different approaches are useful in different circumstances. If we examine a broader set of issues, then either the economic dependency ratio or NTA support ratio may offer useful alternative answers.

Regarding the second question, a number of issues need to be followed up carefully. One issue is about human capital formation. Instead of assuming that it is accumulated by schooling only, one is interested to know by how much the results will change if we assume, as in Ben-Porath (1967) and Manuelli et al. (2012), that human capital may (a) also be accumulated through on-the-job training and (b) depreciate over time.

Finally, we want to close this paper with a reference to public policy. A key message of this paper is that demographic changes, by themselves, may not necessarily be problematic if people (and perhaps the government too) anticipate them appropriately in formulating their responses. However, people sometimes are not able to respond optimally to demographic changes, if existing policies are restrictive and/or provide inappropriate incentives. Another interesting topic for future study is to examine the impact on economic welfare of (a) regulations (or social convention) which restrict age of retirement or (b) pension policies which

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

provide workers financial incentive to retire earlier. We leave these questions to future study.

NOTES

1. The total dependency ratio is usually defined as the sum of the number of persons under a particular (young) age (e.g., 20) and those over another (old) age (e.g., 65), divided by the population between these two ages. The cutoff points of 20 and 65 have been used by the United Nations (UN). Other cutoff points are also commonly used. For example, the UN (http://esa.un.org/unpd/wpp/Excel-Data/population.htm) publishes 5 dependency ratio series with various combination of cutoff points, and the US Census Bureau (https://www.census.gov/population/projections/files/summary/NP2014-T6.xls) uses 18 and 65 to define the dependency ratio. Sanderson and Scherbov (2015) provide a brief history of dependency ratios.

2. A description of the sources of data used in this paper is given in Appendix A.

3. One component—the old-age dependency ratio—increases rapidly from 0.22 in 2010 to a projected level of 0.37 in 2030, and then rises gradually to 0.42 in 2060. Another component—the child dependency ratio—remains more or less unchanged during this period.

4. According to Sanderson and Scherbov (2010, p. 1287), POADR is defined as "the number of people in age groups with life expectancies of 15 or fewer years divided by the number of people at least 20 years old in age groups with life expectancies greater than 15 years."

5. The various conventional dependency ratio series (without immigration) from 2010 to 2060 are also shown in Figure 1.

6. On the other hand, the concept of economic dependency ratio incorporates both extensive and intensive margins of labor force participation. It is defined as "the number of persons in the total population (including the Armed Forces overseas and children) who are not in the labor force, per 100 of those who are in the labor force." (U.S. Bureau of Labor Statistics, 2006, p. 37.) However, while it is easy to obtain the value of this statistic ex post, a lot of judgments about age-specific labor force participation rates in the coming decades need to be made for forecasting. If one wants an objective and transparent way to forecast economic resource dependency in the future, this definition may be less appropriate. On the other hand, our proposed definition aims to keep a major aspect of the conventional dependency ratio, namely the underlying age-specific labor force participation rates are either 0 or 1, unchanged. To distinguish our definition from the economic dependency ratio, we call it EDDR.

7. For the theoretical analysis in this subsection, variables are expressed in terms of adult age, instead of actual age, unless stated explicitly.

8. From now on, the dependence of the optimal choices of years of schooling and age of retirement on cohort b is expressed explicitly.

9. On the other hand, a different prediction regarding the effect of a change in ξ_b arises if one uses the disutility of labor specification. In Section 2.3, we will discuss the difference between the two specifications and explain why we choose the utility jump at retirement specification for the definition of EDDR.

10. In either (14) of Bloom et al. (2014) or (9) of Sánchez-Romero et al. (2016), the marginal cost of delaying retirement, given by the respective right-hand side term, is expressed in level of disutility. Note that the disutility of labor term in Bloom et al. (2014) is measured at the time when the agent retires, but it is discounted back to age 0 in Sánchez-Romero et al. (2016).

11. As seen from (A7a), the marginal benefit of delaying retirement can be measured by the effects on lifetime utility through changes of the consumption path from age 0 to R_b^*, and from age R_b^* to $\Omega - M$. After substituting into (A5), the derivative of lifetime budget constraint equation with respect to age of retirement, these effects of a delay in retirement on lifetime resource can be expressed as

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

the difference of a term related to a change in lifetime labor income and a "consumption drop at retirement" term.

12. This prediction holds for $\sigma < 1$ but disappears if $\sigma = 1$.

13. As seen from the above analysis, two key assumptions—utility jump in retirement and the power utility function—lead to the above result. Under the CIES utility function (with the utility level given by consumption raised to a power depending on σ), the right-hand side of (8), which captures the difference in utility level, is expressed in the unit of $c_b(0, S_b^*, R_b^*)^{1-\frac{1}{\sigma}}$. On the left-hand side of (8), we observe that the marginal utility, which is expressed in the unit of $c_b(0, S_b^*, R_b^*)^{-\frac{1}{\sigma}}$, is present. Moreover, we find that consumption at age 0 varies linearly with wage income $\xi_b h\left(S_b^*\right)$; see (6). As a result, the two terms on the left-hand side of (8)—the product of marginal utility and wage income, as well as the product of marginal utility and a drop in consumption level—are also expressed in the same unit of $c_b(0, S_b^*, R_b^*)^{1-\frac{1}{\sigma}}$. After simplification, it is further found that $(\xi_b)^{1-\frac{1}{\sigma}}$ appears in both sides of (8); thus, ξ_b disappears in (8a).

14. We want to focus on the path of dependency ratio in the coming decades, and the reasons of choosing 2010 to 2060 are as follows. As seen in Appendix A, the birth data are available from 1900 to 2060, and the survival probabilities are available from 1900 to 2000. We use the Lee and Carter (1992) method to extend the survival data to 2060. On the other hand, since we assume the maximum biological age (Ω) is 110, we choose 2010 (i.e., 1900 + 110) as the starting year for the analysis of the dependency ratio. Note also that we simplify by considering the survival data up to age 110 (instead of 120 in the original data set) because the survival probabilities for persons aged above 110 are very close to 0.

15. We assume that the survival probabilities of different cohorts are determined exogenously, following similar approach in many papers in the literature, such as Bloom et al. (2007), d'Albis et al. (2012), and Cervellati and Sunde (2013). This assumption is also consistent with the conventional definition of dependency ratios and enables us to focus on the roles of cohort-invariant versus cohort-varying years of schooling and retirement age. A consequence of this assumption is that the possible feedback from schooling to survival probabilities is ignored. Recent evidence suggests that these differences among various socioeconomic groups have increased in many countries. It is interesting to investigate this feedback effect in future work, which requires modifying the basic life-cycle model to incorporate more sources of heterogeneity. We are grateful to a referee for suggesting the study of the feedback from schooling to survival probabilities.

16. This is chosen to be the same as the normal retirement age of the 1950 cohort published by the US Social Security Administration (http://www.ssa.gov/retire2/retirechart.htm). According to the official definition, the normal retirement age (or full retirement age) is the age at which a person may first become entitled to full or unreduced retirement benefits.

17. This value of percentage drop in consumption at retirement is similar to those commonly used in the literature. Manuelli et al. (2012) use 15%. Galama et al. (2013) cite evidence of 15% to 20% and use 15% in their baseline model. Rogerson and Wallenius (2016) put limits of 10% and 15% and use 10% in their baseline model. This range of values is based on earlier findings, such as Bernheim et al. (2001) and Laitner and Silverman (2005). We also use 10% and 20% in the sensitivity analysis in Section 5.

18. We are grateful to Diego Restuccia for sending the data set. The data set starts from 1876, but we only use the data starting from 1900. Note also that we implicitly assume in (22) that children start going to school at age 6.

19. The increase in optimal age of retirement in Figure 3 appears to be inconsistent with the decline of labor force participation rate of US men, especially for older men, throughout most of the 20th century [see, e.g., Costa (1998) and Maestas and Zissimopoulos (2010)]. Researchers such as Gruber and Wise (1998) suggest that the generous benefit of the social security system may be an important contributing factor. However, we do not include extra factors such as social security in our proposed definition of EDDR because of two reasons. First, social security is not included in the definition of conventional dependency ratio, and we do the same for EDDR to have an appropriate comparison. Second, our focus is economic dependency from 2010 to 2060. According to Maestas and

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

Zissimopoulos (2010, Figure 4), the decreasing trend of retirement age seems to reverse around the 1990s. Not including the social security benefit is probably responsible for the inability to explain the trend of retirement for the earlier cohorts, but its consequence may be relatively minor in the coming decades, which is the period that we focus on.

20. A further investigation of Figure 5B suggests that N_{PoR} is above trend from 2020 to 2045, whereas N_{WA} is slightly below trend during this period. As a result, $EDDR_{PoR}$ is above trend in this period.

21. In 2060, there is a severe under-projection gap of about 7 percentage points in DR_C, but the over-projection gap of 11 percentage points in DR_{OA} is even more severe. Combining them leads to the 4 percentage points over-projection gap for the conventional dependency ratio.

22. One motivation of this decomposition is the perception that while the level of the conventional dependency ratio depends on the cutoff points, its speed of change over a particular period may not be too sensitive to the cutoff points, provided that these points are chosen appropriately. Note that the distinction between the level and speed of change of population aging has also been mentioned in, for example, the study of "...future levels of indicators of ageing and the speed at which they change." [Lutz et al. (2008, p. 716)]

23. The growth rate of a ratio can be expressed as the difference of the growth rate of the numerator versus that of the denominator. The advantage of analyzing the growth rates of various component terms can be seen in Figure 7C and D.

24. The annualized growth rate (expressed in percentage) of a variable x from year 2010 to year t is given by

$$AGR_x(t) = \left[\frac{\ln x\,(t) - \ln x(2010)}{t - 2010}\right](100\%).$$

25. An examination of (24), together with the rather smooth changes in survival functions shown in Figure 2A and the more widely fluctuating birth data in Figure 4, suggest that the hump-shaped pattern is mainly driven by the entering of old-age (at 65) of various baby boom cohorts (1946–1964). The birth numbers of these cohorts clearly exhibit the shape of a hump.

26. It is also observed that the growth rate of $EDDR_{PoR}$ is negative for a few years from 2011. However, this feature does not have any major long-term consequence.

27. These numbers can easily be found from Table 2. For example, adding up –0.0255 and 0.1301 in the "Diff." row for the Baseline Case gives 0.105, after rounding.

28. As in the baseline case, they are chosen to be the same as the normal retirement ages of the corresponding cohorts published by the US Social Security Administration (http://www.ssa.gov/retire2/retirechart.htm).

29. Estimates of intertemporal elasticity of substitution (σ) vary in the literature, and our estimates are consistent with many of them. For example, Attanasio and Weber (1993) estimate that σ is close to 0.8, using U.K. cohort data. Based on existing evidence of the labor supply elasticity, Chetty (2006) argues that there is a tight upper bound of 2 on the coefficient of relative risk aversion, which implies a lower bound of 0.5 for the intertemporal elasticity of substitution in models with time-separable utility. See Attanasio and Weber (2010) for more discussion on the estimates of this parameter.

30. The ranges of these two variables in 1900 and 1950 are even narrower, as seen in Table 2. Note also that years 1990 (for R^*) and 2040 (for S^*) are chosen because the relevant variables in years after these will not affect *EDDR* from 2010 to 2060. Note also that many cells in Table 2 are not filled to avoid repetitive information.

31. The flatness (and around 0) of these paths of the annualized growth rates of $EDDR_{PrW}$ can be further explained by the almost identical annualized growth rates of the pre-working and working populations after 2025, which are plotted in Figure 9C.

32. We have also experimented with $\rho = 1\%$ and $\rho = 3\%$, besides Case 3 ($\rho = 1.5\%$) and Case 4 ($\rho = 2.5\%$). The computed values of *EDDR* for these cases deviate even more from its baseline value when $\rho = 2\%$. For example, when ρ decreases to 1%, the corresponding $DR - EDDR$ decreases from 0.091 to 0.063, which is a larger 31% reduction. And when ρ increases to 3%, the corresponding

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

$DR - EDDR$ rises from 0.091 to 0.139, which is an increase of even larger magnitude (53%). Since our major conclusions (specifically, the conclusion based on eight cases ignoring changes in ρ) do not depend on the values of ρ in the sensitivity analysis, we present the results for $\rho = 1.5\%$ and $\rho = 2.5\%$ in the main text to show that the results are quite sensitive even for the smaller changes in ρ.

33. There are economic and demographic elements in the baseline model. Since the key difference between the conventional and proposed definitions of dependency ratio is mainly due to the underlying economic behavior, we perform sensitivity analysis on economic, but not demographic, parameters. For broader questions such as whether resource dependence in the coming decades will be affected by demographic and/or policy changes, it is interesting to also focus on demographic elements. We leave these issues to future work.

34. People's possible response to population aging by adjusting schooling and retirement choices has also been pointed out in a recent paper by Vogel et al. (2017). However, they focus on the welfare loss of population aging, whereas we focus on the over-projection of resource dependency by the conventional dependency ratio.

35. One can interpret that both the forward-looking demographic approach [e.g., Sanderson and Scherbov (2010)] and EDDR improve over the cohort-invariant cutoff points of conventional dependency ratio. Between these two modified definitions, the EDDR can further be regarded as an improvement over the forward-looking demographic approach by incorporating behavioral response, rather than simply using a statistical formula based on remaining life expectancy.

36. For the reasons just elaborated, we do not consider the intensive margin of labor supply in this paper, but we are aware that any significant change in this aspect may also affect the economic resource dependency. One may extend the baseline model of this paper to examine the impact of demographic changes on both extensive and intensive margins of lifetime labor supply. In particular, it is interesting to examine whether increasing expected wealth of future cohorts, either due to productivity increase or longer life expectancy, may systematically affect their desirable (weekly or annual) work hours during the working phase.

37. The results support our earlier conjecture that the conventional dependency ratio may be quite misleading in measuring the impact of population aging when its pace is rapid.

REFERENCES

Attanasio, O. P. and G. Weber (1993) Consumption growth, the interest rate and aggregation. *Review of Economic Studies* 60, 631–649.

Attanasio, O. P. and G. Weber (2010) Consumption and saving: Models of intertemporal allocation and their implications for public policy. *Journal of Economic Literature* 48, 693–751.

Barro, R. J., N. G. Mankiw and X. Sala-i-Martin (1995) Capital mobility in neoclassical models of growth. *American Economic Review* 85, 103–115.

Ben-Porath, Y. (1967) The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75, 352–365.

Bernheim, B. D., J. Skinner and S. Weinberg (2001) What accounts for the variation in retirement wealth among U.S. households? *American Economic Review* 91, 832–857.

Bloom, D. E., D. Canning, R. K. Mansfield and M. Moore (2007) Demographic change, social security systems, and savings. *Journal of Monetary Economics* 54, 92–114.

Bloom, D. E., D. Canning and M. Moore (2014) Optimal retirement with increasing longevity. *Scandinavian Journal of Economics* 116, 838–858.

Booth, H., R. J. Hyndman, L. Tickle and P. de Jong (2006) Lee-Carter mortality forecasting: A multi-country comparison of variants and extensions. *Demographic Research*, 15, 289–310.

Boucekkine, R., D. de la Croix and O. Licandro (2003) Early mortality declines at the dawn of modern growth. *Scandinavian Journal of Economics* 105, 401–418.

Brown, R. L., J. Durbin and J. M. Evans (1975) Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B (Methodological)* 37, 149–192.

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

Cai, Z. and S.-H. P. Lau (2017) Impact of mortality reductions on years of schooling and expected lifetime labor supply. *Journal of Mathematical Economics* 72, 134–144.
Cervellati, M. and U. Sunde (2013) Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. *Econometrica* 81, 2055–2086.
Chetty, R. (2006) A new method of estimating risk aversion. *American Economic Review* 96, 1821–1834.
Clark, T. E. and K. D. West (2007) Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138(1), 291–311.
Costa, D. L. (1998) *The Evolution of Retirement*. Chicago: University of Chicago Press.
d'Albis, H., S.-H. P. Lau and M. Sánchez-Romero (2012) Mortality transition and differential incentives for early retirement. *Journal of Economic Theory* 147, 261–283.
Galama, T., A. Kapteyn, R. Fonseca and P.-C. Michaud (2013) A health production model with endogenous retirement. *Health Economics* 22, 883–902.
Goldin, C. and L. F. Katz (2008) *The Race Between Education and Technology*. Cambridge, MA: Harvard University Press.
Gruber, J. and Wise, D. A. (1998) Social security and retirement: An international comparison. *American Economic Review Papers and Proceedings* 88(2), 158–163.
Hazan, M. (2009) Longevity and lifetime labor supply: Evidence and implications. *Econometrica* 77, 1829–1863.
Kalemli-Ozcan, S., H. E. Ryder and D. N. Weil (2000) Mortality decline, human capital investment, and economic growth. *Journal of Development Economics* 62, 1–23.
Laitner, J. and D. Silverman (2005) Estimating Life-cycle Parameters from Consumption Behavior at Retirement. National Bureau of Economic Research, Inc., NBER Working Paper: No. 11163.
Lee, R. and A. Mason (2011) *Population Aging and the Generational Economy: A Global Perspective*. Cheltenham, UK: Edward Elgar Publishing Limited.
Lee, R. and A. Mason (2013) Reformulating the Support Ratio to Reflect Asset Income and Transfers. Working Paper (Extended abstract for the Annual meeting of the Population Association of America, 2013).
Lee, R. and T. Miller (2001) Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* 38(4), 537–549.
Lee, R. D. and L. R. Carter (1992) Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87, 659–671.
Loichinger, E., B. Hammer, A. Prskawetz, M. Freiberger and J. Sambt (2017) Quantifying economic dependency. *European Journal of Population* 33, 351–380.
Lutz, W., W. Sanderson and S. Scherbov (2008) The coming acceleration of global population ageing. *Nature* 451 (7 February 2008), 716–719.
Maestas, N. and J. Zissimopoulos (2010) How longer work lives ease the crunch of population aging. *Journal of Economic Perspectives* 24(1), 139–160.
Manuelli, R. E., A. Seshadri and Y. Shin (2012) Lifetime Labor Supply and Human Capital Investment. Working Paper (Washington University St. Louis).
Rapach, D. E., J. K. Strauss and M. E. Wohar (2008) Forecasting stock return volatility in the presence of structural breaks. In: D. E. Rapach and M. E. Wohar (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, pp. 381–416. Bingley, UK: Emerald Group Publishing Limited.
Rogerson, R. and J. Wallenius (2016) Retirement, home production and labor supply elasticities. *Journal of Monetary Economics* 78, 23–34.
Sánchez-Romero, M., H. d'Albis and A. Prskawetz (2016) Education, lifetime labor supply, and longevity improvements. *Journal of Economic Dynamics and Control* 73, 118–141.
Sanderson, W. C. and S. Scherbov (2005) Average remaining lifetimes can increase as human populations age. *Nature* 435 (9 June 2005), 811–813.
Sanderson, W. C. and S. Scherbov (2010) Remeasuring aging. *Science* 329 (10 September 2010), 1287–1288.
Sanderson, W. C. and S. Scherbov (2015) Are we overly dependent on conventional dependency ratios? *Population and Development Review* 41(4), 687–708.

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

Tuljapurkar, S., N. Li and C. Boe (2000) A universal pattern of mortality decline in the G7 countries. *Nature*, 405, 789-792.
United Nations (2013) Probabilistic Population Projections based on the World Population Prospects: The 2012 Revision. http://esa.un.org/unpd/wpp/Excel-Data/population.htm
U.S. Bureau of Labor Statistics (2006) A new look at long-term labor force projections to 2050. *Monthly Labor Review*, November, 19–39.
U.S. Census Bureau (2003) 2003 Statistical Abstract: Mini-Historical Statistics, No. HS-13. https://www.census.gov/library/publications/2003/compendia/statab/123ed/hist.html
U.S. Census Bureau (2014) 2014 National Population Projections. https://www.census.gov/population/projections/files/summary/NP2014-T1.xls
U.S. National Center for Health Statistics (2015a) National Vital Statistics Reports, Vol. 64, No. 1. http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_01.pdf
U.S. National Center for Health Statistics (2015b) National Vital Statistics Reports, Vol. 64, No. 6. http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_06.pdf
Vogel, E., A. Ludwig and A. Börsch-Supan (2017) Aging and pension reform: Extending the retirement age and human capital formation. *Journal of Pension Economics and Finance* 16, 81–107.
Yaari, M. E. (1965) Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32, 137–150.

APPENDIX

Appendix A gives the sources of data used in this study, Appendix B presents detailed derivation of the first-order conditions of the model, and Appendix C describes briefly the Lee–Carter method.

APPENDIX A

Data source:

(A) Birth data (males and females combined): 1900 to 2060.

1901 to 1908: Own calculation, based on (a) infant death data from U.S. Census Bureau (2003) and (b) assuming 5% infant mortality rate.

1909 to 2001: U.S. Census Bureau (2003).

2002 to 2013: U.S. National Center for Health Statistics (2015a).

2014: U.S. National Center for Health Statistics (2015b).

2015 to 2060: U.S. Census Bureau (2014).

(B) Survival probabilities (males and females combined): 1900 to 2000

Berkeley Mortality Database (http://www.demog.berkeley.edu/~bmd/).

(C) Schooling data (males and females combined): 1900 to 1975

Goldin and Katz (2008): Including black and white individuals.

(D) Dependency ratios: 2010 to 2060.

United Nations (2013).

APPENDIX B

We solve the problem in two steps. The first step is to obtain the individual's optimal consumption path, conditional on years of schooling, and age of retirement. The objective function (2) is equivalent to

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

$$\int_0^{\Omega-M} e^{-\rho x} l\,(x)$$

$$\left\{[I_S(x)+I_W(x)]\left[\frac{c(x)^{1-\frac{1}{\sigma}}-1}{1-\frac{1}{\sigma}}\right]+[1-I_S(x)-I_W(x)]\right.$$
$$\left.\left[\frac{[(1+\theta)\,c(x)]^{1-\frac{1}{\sigma}}-1}{1-\frac{1}{\sigma}}\right]\right\}dx,$$

and the flow budget constraint (3) can be written as

$$a'\,(x)=[r+\mu\,(x)]\,a\,(x)+I_W(x)h\,(S)-c\,(x)\,,$$

where $I_S(x)$ is an indicator function that takes the value of 1 if $x \le S$ and 0 otherwise, and $I_W(x)$ is an indicator function that takes the value of 1 if $S < x \le R$ and 0 otherwise (Note that these indicator functions, as well as those in (9), (10), and (11), are defined to indicate the individual's various life-cycle stages. However, the indicator functions in (9) to (11) depend on both time (t) and year of birth (b) to facilitate cross-sectional summation).

It can be shown that the necessary first-order conditions for this dynamic optimization problem are

$$l(x)\left\{[I_S(x)+I_W(x)]\,c(x)^{-\frac{1}{\sigma}}+[1-I_S(x)-I_W(x)]\,(1+\theta)^{1-\frac{1}{\sigma}}\,c(x)^{-\frac{1}{\sigma}}\right\}=\eta(x),$$

and

$$\eta(x)\,[r+\mu\,(x)]=\rho\eta(x)-\eta'(x),$$

where $\eta(x)$ is the co-state variable associated with the state variable.

Solving these sets of equations, the optimal consumption path is characterized by:

(a) For $x < R$ or $x > R$,

$$c'(x)=\sigma(r-\rho)c(x). \tag{A1}$$

(b) Because of

$$\lim_{x\to R^-}\eta\,(x)=\lim_{x\to R^+}\eta\,(x)\,,$$

it can be shown that there is a discontinuous drop in consumption level at retirement, with

$$\lim_{x\to R^+}c(x)=(1+\theta)^{\sigma-1}\lim_{x\to R^-}c\,(x)\,. \tag{A2}$$

Combining (A1) and (A2) gives (5) in the main text.

The intertemporal budget constraint at age 0 is given by

$$\int_0^R e^{-rx}l_b\,(x)\,c_b(x,S,R)dx+\int_R^{\Omega-M} e^{-rx}l_b\,(x)\,c_b(x,S,R)dx$$
$$=\xi_b h(S)\int_S^R e^{-rx}l_b\,(x)\,dx, \tag{A3}$$

where the dependence of the consumption path on S and R is written explicitly. Substituting (5) into (A3) leads to (6).

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

Differentiating (A3) with respect to S and R, respectively, we have

$$\int_0^R e^{-rx} l_b(x) \frac{\partial c_b(x,S,R)}{\partial S} dx + \int_R^{\Omega-M} e^{-rx} l_b(x) \frac{\partial c_b(x,S,R)}{\partial S} dx$$

$$= \xi_b h'(S) \int_S^R e^{-rx} l_b(x)\, dx - e^{-rS} l_b(S)\, \xi_b h(S), \tag{A4}$$

and

$$\int_0^R e^{-rx} l_b(x) \frac{\partial c_b(x,S,R)}{\partial R} dx + \int_R^{\Omega-M} e^{-rx} l_b(x) \frac{\partial c_b(x,S,R)}{\partial R} dx$$

$$+ e^{-rR} l_b(R) \left[\lim_{x \to R^-} c_b(x,S,R) - \lim_{x \to R^+} c_b(x,S,R) \right] = e^{-rR} l_b(R)\, \xi_b h(S). \tag{A5}$$

Conditional on the optimal consumption path (5), the second step is to obtain the first-order conditions for optimal years of schooling and age of retirement. After substituting (5) into the objective function (2), we define the objective function as a function of S and R:

$$V(S,R) = \int_0^R e^{-\rho x} l_b(x) \left[\frac{[c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] dx$$

$$+ \int_R^{\Omega-M} e^{-\rho x} l_b(x) \left[\frac{[(1+\theta)\, c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] dx. \tag{A6}$$

Differentiating (A6) with respect to S and using (5) and (A4), we obtain

$$\frac{\partial V(S,R)}{\partial S} = \int_0^R e^{-\rho x} l_b(x)\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial S} dx$$

$$+ \int_R^{\Omega-M} e^{-\rho x} l_b(x)\, (1+\theta)^{1-\frac{1}{\sigma}}\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial S} dx$$

$$= c_b(0,S,R)^{-\frac{1}{\sigma}} \xi_b \left[h'(S) \int_S^R e^{-rx} l_b(x)\, dx - e^{-rS} l_b(S)\, h(S) \right]$$

Therefore, the first-order condition for S $\left(\frac{\partial V(S,R)}{\partial S} = 0\right)$ is given by (7).

Differentiating (A6) with respect to R, we have

$$\frac{\partial V(S,R)}{\partial R} = \int_0^R e^{-\rho x} l_b(x)\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial R} dx$$

$$+ \int_R^{\Omega-M} e^{-\rho x} l_b(x)\, (1+\theta)^{1-\frac{1}{\sigma}}\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial R} dx$$

$$- e^{-\rho R} l_b(R) \left\{ \lim_{x \to R^+} \left[\frac{[(1+\theta)\, c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] \right.$$

$$\left. - \lim_{x \to R^-} \left[\frac{[c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] \right\}$$

$$= MB_R - MC_R, \tag{A7}$$

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

where MB_R and MC_R are given, after using (5) and (A5) to simplify, as follows:

$$MB_R = \int_0^R e^{-\rho x} l_b\,(x)\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial R} dx$$

$$+ \int_R^{\Omega - M} e^{-\rho x} l_b\,(x)\,(1+\theta)^{1-\frac{1}{\sigma}}\, c_b(x,S,R)^{-\frac{1}{\sigma}} \frac{\partial c_b(x,S,R)}{\partial R} dx$$

$$= c_b(0,S,R)^{-\frac{1}{\sigma}} \left[\int_0^R e^{-rx} l_b\,(x)\, \frac{\partial c_b(x,S,R)}{\partial R} dx + \int_R^{\Omega - M} e^{-rx} l_b\,(x)\, \frac{\partial c_b(x,S,R)}{\partial R} dx \right]$$

$$= c_b(0,S,R)^{-\frac{1}{\sigma}} e^{-rR} l_b\,(R) \left\{ \xi_b h\,(S) - \left[\lim_{x \to R^-} c_b(x,S,R) - \lim_{x \to R^+} c_b(x,S,R) \right] \right\}$$

$$= c_b(0,S,R)^{-\frac{1}{\sigma}} e^{-rR} l_b\,(R) \left\{ \xi_b h\,(S) - \left[1 - (1+\theta)^{(\sigma-1)} \right] \lim_{x \to R^-} c_b(x,S,R) \right\}, \qquad \textbf{(A7a)}$$

and

$$MC_R = e^{-\rho R} l_b\,(R) \left\{ \lim_{x \to R^+} \left[\frac{[(1+\theta)\, c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] - \lim_{x \to R^-} \left[\frac{[c_b(x,S,R)]^{1-\frac{1}{\sigma}} - 1}{1-\frac{1}{\sigma}} \right] \right\}$$

$$= e^{-\rho R} l_b\,(R) \left[(1+\theta)^{(\sigma-1)} - 1 \right] \frac{[\lim_{x \to R^-} c_b(x,S,R)]^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}. \qquad \textbf{(A7b)}$$

Therefore, the first-order condition for R $\left(\frac{\partial V(S,R)}{\partial R} = 0\right)$ is given by (8). ■

APPENDIX C

The standard Lee–Carter model is specified as follows:

$$\ln\left(m_{x,t}\right) = a_x + b_x k_t + \varepsilon_{x,t}, \qquad \textbf{(A8)}$$

where $m_{x,t}$ ($x = 1, \ldots, \Omega$; $t = 1, \ldots, T$) is the mortality rate at age x in year t; a_x is a set of age-specific constants reflecting the general pattern of mortality by age; k_t is a time-varying index of the general level of mortality; b_x is a set of age-specific constants describing the responsiveness of mortality at age x to variations in the general level of mortality; and $\varepsilon_{x,t}$ is the error term at age x in year t, which is usually assumed to be homoscedastic over time. To obtain a unique solution, the Lee and Carter (1992) impose constraints that $\sum_x b_x = 1$ and $\sum_t k_t = 0$.

We apply the Lee–Carter estimation procedures to annual life tables for the USA by year of birth from 1900 to 2000. The age-specific constants a_x are obtained by taking the average of $\ln\left(m_{x,t}\right)$ over time: $\widehat{a}_x = \frac{1}{T} \sum_t \ln\left(m_{x,t}\right)$, where $T = 101$. The method of singular value decomposition (SVD) is then used to estimate b_x and k_t. The SVD decomposes the difference of matrix $\ln\left(m_{x,t}\right)$ and $\widehat{a}_x$ into the product of three matrices, consisting of the age component, the singular values, and the time component (Let $\mathbf{M}$ be the $\Omega \times T$ matrix such that its (x,t)-th element is $\ln\left(m_{x,t}\right) - \widehat{a}_x$. The SVD of matrix $\mathbf{M}$ can be written as $\mathbf{M} = \mathbf{USV}'$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, and $\mathbf{S}$ is a diagonal matrix containing singular values of $\mathbf{M}$. Note that matrices $\mathbf{M}$ and $\mathbf{S}$ specified in this Appendix are different from M and S (which are scalars and not in bold) in the main text). The vector of general

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press

level of mortality k_t is obtained by two steps. The crude values of k_t is initially derived from the first vector of the time component and the first singular value; in the same step, vector b_x can be derived directly from the first vector of the age component (After normalization, the formulae are given by: $\widehat{b}_x = \frac{u_{x,1}}{\sum_x u_{x,1}}$ and $\widehat{k}_t = s_1 v_{t,1} \sum_x u_{x,1}$, where s_1 is the largest singular value, and $u_{x,1}$ and $v_{t,1}$ are the elements of **U** and **V**). Each crude k_t is further refined by matching the fitted life expectancy at birth with the actual life expectancy for a given year (Lee and Miller, 2001, p. 540). Standard ARIMA time series methods are used to model the refined k_t series (We use the random walk with drift specification, following Lee and Carter (1992) and Booth et al. (2006, Section 2.1). The estimated values of the drift term is –1.4144). Since a_x and b_x are assumed to be time-invariant, the mortality rates and hence the survival probabilities beyond 2000 can be conveniently obtained by forecasting values of k_t in the Lee–Carter model.

https://doi.org/10.1017/S1365100518000962 Published online by Cambridge University Press